I. Remarks

By the enclosed amendment, newly added claim 25 is also pending in the subject application. These claims are believed to be properly included in the examination as set forth.

Claims herein under examination are claims 7-11, 22, and newly added claim 25.

A. Amendments to the claims:

Claims 7 and 11 have been amended; no new matter has been added by this amendment. Support for the amendment to claims 7 and 11 is found throughout the application and particularly at page 10, line 25 - page 11, line 5; page 34, lines 14-16; page 35, lines 8-18; and claim 22.

Claim 25 has been newly added by this amendment; no new matter has been added with these new claims. Support for new claim 25 is found throughout the application and particularly at page 10, line 25 - page 11, line 5; page 34, lines 14-16; page 35, lines 8-18; page 36, lines 8-13; and claims 11, 22.

B. Amendments to the specification:

The disclosure has been objected to because of the following informalities.

(a) The application contained sequence disclosures encompassed by the definitions set forth in 37 C.F.R. § 1.821 (a)(1) and (a)(2) but failed to comply with the requirements set forth in 37 C.F.R. § 1.821- § 1.825. A separate response for the Notice to Comply with Requirements for Patent Applications Containing Nucleotide and/or Amino Acid Sequences has been submitted concurrently with this response to bring the disclosure into full compliance with 37 C.F.R. § 1.821- § 1.825.

Applicants' agent respectfully requests entry of the amendments (1) - (3) to the specification as provided in the section titled "Specification Amendments". These amendments bring the specification into compliance with 37 C.F.R. § 1.821- § 1.825 but do not introduce new matter. The following remarks provide support for each amendment requested; the numbering corresponds to the respective numbering in the Specification Amendments section.

(1) The first amendment adds a comma to provide appropriate sentence structure. The second amendment adds the term "SEQ ID NO: 1" to describe the sequence as required by 37 C.F.R. § 1.821(d). As such, no new matter has been added by these amendments.

(2) The first amendment adds a "t" to provide the correct spelling of the word "streptavidin". This correct spelling is found in the instant specification, for example, at page 36, line 20. The second amendment adds the term "SEQ ID NO: 1" to describe the sequence as required by 37 C.F.R. § 1.821(d). As such, no new matter has been added by these amendments.

(3) This amendment has added the term "SEQ ID NO: 1" to describe the sequence as required by 37 C.F.R. § 1.821(d). As such, no new matter has been added by this amendment.

(b) The specification contained boxes embedded in the text which made it unclear the characters to which the boxes referred. An example was cited on page 45, line 15. Applicant's agent respectfully requests entry of the amendments (4) - (6) to the specification provided in the section titled "Specification Amendments". The following remarks provide support for each amendment requested; the numbering corresponds to the respective numbering in the Specification Amendments section.

(4) This amendment removes the objected to box character and replaces it with the Greek letter alpha, which is the letter utilized according to scientific convention to indicate a portion of a Class I MHC molecule. An excerpt from an immunobiology book is enclosed to demonstrate this convention. See the legend of Figure 4.3 from Charles A. Janeway, Jr. et al., Immunobiology: the Immune System in Health and Disease 118-199 (Current Biology Publications 4th ed. 1999). The second amendment removes the objected to box character and replaces it with the Greek letter β. This letter is utilized according to scientific convention to indicate a portion of a Class I MHC molecule. Support for this convention can also be found in the enclosed excerpt from Immunobiology: the Immune System in Health and Disease (see the legend of Figure 4.3) and in the instant specification at, for example, page 11, line 18. As such, no new matter has been added by these amendments.

(5) The first amendment corrects a grammatical error by replacing "at" with "are". The second amendment deletes the objected to box characters and replaces them with the Greek letters alpha and beta, respectively. Support for this amendment is found throughout the instant specification and in particular at page 57, lines 16-18, where the characters that the boxes denote are specified. As such, no new matter has been added by these amendments.

(6) This amendment deletes the objected to box character and replaces it with the Greek letter β. Support for this amendment is found throughout the instant specification and in particular at page 56, lines 25-31. As such, no new matter has been added by this amendment.

II. Claim rejections under 35 U.S.C. § 112

35 U.S.C. § 12, First Paragraph:

1) Claims 7-10 and 22 stand rejected under 35 U.S.C. § 112, first paragraph, as allegedly failing to comply with the written description requirement. It is alleged that the claim(s) contain subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

Specifically, the Office has alleged that the limitation in claim 7, which recites "a self-assembling fusion polypeptide wherein said fusion polypeptide (i) is capable of forming a stable homomultimer", is not supported by the specification or claims as originally filed. The Office alleges that this feature constitutes new matter. Applicants respectfully traverse.

Contrary to the Office's assertion, this feature of claims 7-10 and 22 is fully supported in the specification. The instant specification teaches that the claimed fusion polypeptide comprises a T cell antigen presenting domain fused to an oligomerization domain. (See page 36, lines 1-3 and original claim 4). The word "fused" indicates that these two domains are connected and are therefore a single polypeptide. The definition of "fusion polypeptide" provided in the instant specification provides for this construction:

> "A 'fusion polypeptide' refers to a chimeric polypeptide molecule comprised of two or more polypeptides joined together by peptide bonds wherein the polypeptides do not naturally occur in such a configuration." (Instant specification at page 17, lines 4-6)

The original specification also teaches that the claimed fusion polypeptides of the instant invention can form multimers spontaneously. (See page 36, lines 8-11.) The embodiment of the instant invention presented in claim 7 (and those dependent therefrom) provides that these fusion polypeptides form multimers spontaneously through the self-assembly of the oligomerization domain. The specification teaches this embodiment at page 36, lines 8-14:

> "The fusion polypeptides of the invention can form multimers spontaneously, without the need for chemical modifications, by self-assembly of the oligomerization domain or by binding of the an oligomerization domain to a second platform molecule. For example, if the oligomerization domain is a self-assembling molecule such as a leucine zipper, the fusion polypeptides will spontaneously form dimers in a solution with an appropriate salt concentration and pH. If the oligomerization domain is a polypeptide factor that forms homotetramers, the fusion polypeptide will form tetramers." (emphasis added)

As demonstrated by the above text, the oligomerization domain of the fusion polypeptide is the portion of the polypeptide that confers the ability to self-assemble upon it. However, the entire polypeptide forms a homomultimer as a result of the oligomerization domain because it is a single polypeptide molecule. Accordingly, Applicants assert that the rejected limitation in claims 7-10 and 22 does not represent new matter. The original specification and claims fully support the feature. Withdrawal of this rejection is therefore respectfully requested.

2) Claims 7-11 and 22 stand rejected under 35 U.S.C. § 112, first paragraph, as allegedly failing to comply with the written description requirement. It is alleged that the claim(s) contain subject matter that was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

Specifically, the Office has alleged that the genus of self-assembling peptides is large and that Applicants have only disclosed a limited number of self-assembling polypeptides. The Office further alleges that the instant specification discloses 1) only MHC molecules as the antigen presenting portion of the fusion construct and 2) only discloses leucine zippers as the "self-assembling" oligomerization domains. As such, the Office alleges that Applicants were not in possession of the claimed invention as claimed. Applicants respectfully traverse.

Claim 22 claims the embodiments of the instant invention where the T cell antigen presenting domain comprises that of an MHC molecule and where the oligomerization domain comprises a leucine zipper. Therefore, the rejection set forth above is moot with respect to claim 22.

Applicants have amended claims 7 and 11 to specify that the T cell antigen presenting domain comprises that of an MHC molecule. Therefore, with respect to the antigen presenting domain, this rejection is moot.

With respect to the remaining portion of the rejection of claims 7-11, Applicants were in possession of the claimed invention at the time of filing. As required by written description, the instant specification describes the claimed invention with sufficient identifying characteristics so that a skilled artisan would recognize possession by Applicants. The claimed invention comprises a T cell antigen presenting domain fused to an oligomerization domain wherein the oligomerization domain self-assembles to form homomultimers spontaneously (see page 36, lines 1-7; 8-14). The specification teaches that the oligomerization domain can be self-assembling polypeptides, which form homomultimers. It provides the leucine zipper domain as a non-limiting example (see page 36, lines 1-7).

Self-assembling polypeptides that spontaneously form oligomers were well-known and conventional in the art when the instant application was filed. It is well-recognized in case law that an applicant need not disclose in detail that which is conventional or well-known to one of ordinary skill in the art. Hybritech Inc. v. Monoclonal Antibodies, Inc., 231 USPQ 81, 94 (Fed. Cir. 1986). Further, the specification "need not teach, and preferably omits, what is well known in the art." Spectra-Physics, Inc. v. Coherent, Inc., 3 USPQ 2d 1737, 1743 (Fed. Cir. 1987). Although the specification cites leucine zippers, other such self-assembling polypeptides were well-appreciated in the art.

For example, synthetic peptides that self-assemble into various β-sheet structures, which are polypeptide homomultimer structures, were widely known. See the following: 1) a synthetic peptide homologous to the β protein of brain amyloid in Kirschner et al., *Proc. Natl. Acad. Sci. USA.* (1987), 84(19):6953 at 6953, col. 2, lines 8-30; 2) glutamine repeats as polar zippers in Perutz et al., *Proc. Natl. Acad. Sci. USA.* (1994), 91(12):5355 at p. 5355, at col. 2, lines 31-34 and Figs. 1, 2(b); 3) glutamine repeats make proteins oligomerize in Stott et al., *Proc. Natl. Acad. Sci. USA.* (1995), 92(14):6509 at Abstract and col. 1, lines 10-20; and 4) self-assembly of peptides into β-sheet tapes in Aggeli et al., *Nature.* (1997), 386(6622):259 at Abstract and Fig. 2(b).

Peptides from the Hepatitis B surface antigen (HBsAg) and an HIV type 1 envelope protein were also recognized as self-assembling peptides. See the following: 1) synthetic peptide from HBsAg self-assembles in Manivel et al., *J. Immunol.* (1992), 148(12):4006 at 4007, col. 1, lines 7-12; and 2) self-assembling peptide of HIV type 1 envelope protein in Tripathy et al., *J. Immunol.* (1992), 148(12):4012 at 4013, col.1, lines 29-33. Adenoviral fiber proteins also were known to be capable of self-assembling into trimers. See Novelli et al., *J. Biol. Chem.* (1991), 266(14):9299 at 9302, 1st line of Discussion. Indeed, the general study of self-assembling oligomeric peptides was well-established. See Xiong et al., *Proc. Natl. Acad. Sci. USA.* (1995), 92(14):6349.

Therefore, the prior art demonstrates that self-assembling polypeptides were well-recognized and conventional to those skilled in the art. The instant specification teaches that the oligomerization domain is a self-assembling polypeptide capable of forming homomultimers, pointing to the well-known leucine zipper domain as a non-limiting example. Based on these teachings and the state of the art at the time of the instant invention, Applicants assert that the written description requirement is fulfilled for the claimed invention. There is no doubt that the skilled artisan would have immediately understood that Applicants were in possession of the instant claims 7-11 at the time of filing.

Based on the above arguments, Applicants respectfully request that the written description rejection be withdrawn and claims 7, 11, and 22 be allowed to issue.

35 U.S.C. § 12, Second Paragraph:

Claim 10 stands rejected under 35 U.S.C. § 112, second paragraph, as allegedly failing to particularly point out and distinctly claim the subject matter which Applicants regard as their invention.

The Office has alleged that there is no antecedent basis for the limitation "the polynucleotide of claim 47" in claim 10. Applicants respectfully traverse.

In the Response and Amendment filed by Applicants on March 21, 2003, claim 10 was amended to change the dependency from claim 4 to claim 7. (See the Remarks section of the Response and Amendment as filed on March 21, 2003.) A strikethrough of the "4" was performed to indicate this amendment. This strikethrough is very difficult to see because of the font utilized by Applicants in their response. Therefore, Applicants have again presented claim 10 in amended form using double brackets to indicate the deletion of the "4" in the set of claims enclosed herein. As such, withdrawal of this rejection is respectfully requested.

III. Claim rejections under 35 U.S.C. § 102(b)

Claims 7-10 and 22 stand rejected under 35 U.S.C. § 102(b) as being allegedly anticipated by Scott et al., *J. Exp. Medicine* 183:2087-2095 (1996) referred to hereafter as 'Scott".

The Office has alleged that Scott et al. teaches a polynucleotide which encodes a fusion polypeptide comprising a T cell antigen presenting domain and an oligomerization domain which can form a stable homomultimer. It is also alleged that the reference teaches a gene delivery vehicle vector comprising the polynucleotide, host cells comprising the polynucleotide, the polypeptide expressed from the polynucleotide, and a leucine zipper domain as the oligomerization domain. Applicants respectfully traverse.

The Scott reference does not anticipate the instant invention because it teaches a polynucleotide system encoding for polypeptides which are only capable of forming heterodimers, i.e. where two different polypeptide chains self-assemble. In contrast, the instant invention requires a polynucleotide encoding a single polypeptide only capable of forming a homomultimer, i.e. where two identical polypeptide chains self-assemble. As Scott discloses a fusion polypeptide incapable of forming a homomultimer, the reference cannot anticipate the instant invention as claimed.

The Scott reference examines the effect of enhanced murine MHC IA α and IA β chain pairing on the production of secreted MHC IA^d molecules. To accomplish this, both MHC IA chains (α and β) were modified to include a complementary leucine zipper at their respective COOH terminus. These complementary leucine zippers force the association of the IA α and β chains. (See page 2087, 1st paragraph, last sentence; and 2cd paragraph, third sentence.) Figure 1A illustrates these polynucleotide constructs with the two different MHC IA chains and complementary leucine zippers (see page 2089, Figure 1A under "Modified cDNAs".) Expression of these polynucleotides produces polypeptides that associate to form only the paired, heterodimeric product (i.e. an IA α chain will not bind to another IA α chain.)

> "Since each chain is engineered to have a COOH-terminal peptide of opposite charge, the addition of the zipper produces only the correctly paired heterodimeric product." (Scott et al., page 2089, 1st full paragraph, third sentence.)

Therefore, the polynucleotides taught by the Scott reference does not and cannot produce a homomultimer as required in claim 7, step (ii). The Office acknowledges that the polypeptides produced in Scott are heterodimers, not homodimers:

> "...the teachings of Scott regarding the manufacture of *heterodimers* of promiscuous MHC domains using linking domains..." (emphasis added, Paper 14, page 6, section 5, 4th paragraph, 1st sentence.)

The claimed invention requires a polynucleotide encoding for a fusion polypeptide capable of forming homomultimers by self-assembly via the oligomerization domain. Therefore, the Scott reference does not anticipate the instant invention. Withdrawal of this rejection is respectfully requested.

IV. Claim rejections under 35 U.S.C. § 103(a)

Claim 11 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Scott et al., *J. Exp. Medicine* 183:2087-2095 (1996) ["Scott"] in view of U.S. Patent No. 6,015,884 to Schneck et al. ['884].

The Office alleges that it would have been obvious for a person of ordinary skill in the art to combine the teachings of Scott regarding the production of MHC Class II heterodimers with the teachings of the '884 patent regarding the covalent linkage of an MHC peptide to a soluble MHC Class II heterodimer to arrive at the instant invention. Applicants respectfully traverse.

The establishment of a *prima facie* case of obviousness requires, in part, that all claim limitations must be taught or suggested by the references. (MPEP 2142-2143). The instant invention is not obvious

in light of the cited prior art because the cited references fail to provide all of the claim limitations of claim 11.

Claim 11 requires two, separate polynucleotides. In claim 11, the first polynucleotide encodes for a polypeptide comprising a T cell antigen presenting domain fused to an oligomerization domain (step i). The second, separate polynucleotide in claim 11 encodes for a T cell epitope which binds specifically to the T cell antigen presenting domain encoded for in the first polynucleotide (step ii). Neither the Scott reference nor the '884 patent provides the first polynucleotide required by step (i). In addition, neither reference provides the separate, second polynucleotide required in step (ii).

As discussed above, the Scott reference teaches a polynucleotide system encoding for polypeptides which are only capable of forming heterodimers, i.e. where two different polypeptide chains self-assemble. It does not teach the a polynucleotide that encodes for a polypeptide capable of self-assembling into homomultimers. Accordingly, Applicants assert that the Office has failed to provide the polynucleotide required by step (i) of claim 11.

Next, the Scott reference and the '884 patent fail to teach the separate, second polynucleotide encoding for a T cell epitope.The Office has cited the entire patent in general and specifically Figure 1C as teaching the covalent linkage of an MHC peptide to a soluble Class II heterodimer, where the epitope binds specifically in the groove of the MHC Class II molecule. It appears that the Office has extrapolated from this figure and the patent the following: 1) that it teaches a first polynucleotide, encoding for a fusion protein, linked to a second polynucleotide, encoding for an epitope and 2) that, prior to the linkage of the first polynucleotide to the second polynucleotide in a recombinant system, the two polynucleotides existed as separate entities in the recombinant system. Through this extrapolation, the Office allegedly finds the second polynucleotide encoding for a T cell epitope.

Applicants agree that the '884 patent teaches the expression of heterodimeric proteins using a recombinant baculovirus system. However, Applicants cannot find a teaching or a suggestion of a polynucleotide encoding for a T cell epitope as required by the instant invention. The expression vector maps provided in the '884 patent (see Figure 2 in '884) are silent with respect to T cell epitopes. The examples that teach the expression and detection of these soluble heterodimeric proteins are also silent with respect to T cell epitopes (see '884, col. 19 , line19 - col. 22, line 32). T cell epitopes are provided to cells as peptides, not as polynucleotide sequences, throughout the '884 patent and even in Figure 1C cited by the Office (see Figure 1C where the "MCC peptide" is bound to the MHC Class II groove.) As far as Applicants are able to determine, the '884 patent fails to provide a teaching of any polynucleotide sequence that encodes for a T cell epitope let alone a teaching of a polynucleotide that encodes for an

epitope to bind specifically to the T cell antigen presenting domain encoded for in a specific, first polynucleotide.

Accordingly, Applicants assert that the Office has failed to provide the polynucleotide required by step (ii) of claim 11. Without all elements of the instant invention, one skilled in the art cannot be motivated to combine the references in the manner suggested by the Office. Nor, without all of the elements of the instant invention, could one skilled in the art reasonably expect to succeed in making the combination. The absence of all elements precludes the establishment of a prima facie case of obviousness. Applicants assert that the invention cannot be obvious in light of the cited references. Withdrawal of this rejection is respectfully requested.

V. Newly added claim 25

Claim 25 claims a recombinant system comprising two separate polynucleotides. The first polynucleotide encodes a fusion polypeptide wherein the T cell antigen presenting domain comprises that of an MHC molecule and wherein the oligomerization domain comprises a leucine zipper. The second polynucleotide encodes a T cell epitope that binds specifically to the antigen presenting domain of the fusion polypeptide.

Applicants assert that this claim is both novel and non-obvious over the prior art that has been cited by the Office for the reasons discussed above in the individual rejections. Moreover, Applicants note the Office's comments at Paper No. 14, page 4, lines 14-17, which recognize that both the MHC molecule and the leucine zipper portions of the fusion polypeptide are fully disclosed in the instant specification. For these reasons, Applicants respectfully requests that claim 25 be allowed to issue.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 14, Issue of May 15, pp. 9299-9303, 1991
Printed in U.S.A.

Assembly of Adenovirus Type 2 Fiber Synthesized in Cell-free Translation System*

(Received for publication, October 3, 1990)

Armelle Novelli and Pierre A. Boulanger‡

From the Laboratoire de Virologie et Pathogénèse Moléculaires, Faculté de Médecine, Boulevard Henri IV, 34060 Montpellier, France

Physicochemical and functional analyses of the translation products of fiber mRNA in rabbit reticulocyte lysate suggested that fiber polypeptide chains (monomers) were capable of self-assembling *in vitro*, forming trimeric fibers (trimers) without direct intervention of any other adenovirus-coded protein or cell nuclear matrix component. Kinetic studies showed that trimer formation occurred at a rate six times lower than that of fiber polypeptide synthesis. Fiber assembly was found to be relatively inefficient *in vitro*, with only 25–30% fiber polypeptides trimerized after 4-h translation reaction. The rate constant for fiber subunit assembly, extrapolated from the kinetic curves of trimer formation, was found to be in the order of magnitude of 10^6 M^{-1} s^{-1}, with a $t_{1/2}$ of 1.3 h at 30 °C. A latence phase of ~40 min in the appearance of the first detectable trimers indicated that fiber assembly did not occur co-translationally, suggesting the existence of rate-limiting intermediate step(s) during assembly.

The vertex capsomer of the adenovirus (Ad)[1] particle, referred to as the penton, is a complex structure consisting of two elements linked by non-covalent bonds: the penton base, anchored in the capsid where it interacts with five neighboring hexons on the 5-fold symmetry axis of the virion, and its distal projection called the fiber. The penton, penton base, and fiber have all been isolated as free capsomers from adenovirus infected cells and from disrupted virions as well (reviewed in Refs. 1 and 2). The fiber of adenovirus type 2 (Ad2) is a very asymmetrical protein of 582 amino acids with an axial ratio of greater than 15 in velocity gradient centrifugation (3) and is visible under the electron microscope as a rod-like structure of which length has been recently re-evaluated (4, 5), terminated by a spherical knob (1, 2). Based on these morphological properties, the fiber has been arbitrarily divided into three domains, *viz.* tail, shaft, and terminal knob. The polarity of the polypeptide chains in the fiber structure has been unambiguously determined from experiments using partial proteolysis, chemical cleavage, and anti-peptide sera (6–8). All the data are consistent with the N terminus being located in the tail region interacting with the penton base and the C terminus contained in the knob which binds to the cell receptors (9).

In Ad2, the shaft is formed by 22 repeats of a 15-residue motif in a β-sheet conformation (10). The molecular mass of the Ad2 fiber polypeptide unit is 62,000 daltons (62 kDa), whereas the native fiber capsomer behaves as a protein of 170–180 kDa (3, 11, 12). This suggested a stoichiometry of three polypeptide subunits to one fiber capsomer (13, 14). Although a model for a thermodynamically stable fiber comprising two identical polypeptide chains has been proposed (10), more recent crystallographic analyses have strong implications of the fiber being a trimer (5). Ad2 fiber has been shown to be glycosylated (15) and the nature of the carbohydrate residue identified as a very unusual *O*-linked *N*-acetylglucosamine located in the N-terminal third of the fiber polypeptide chain (16).

The aim of the present study was to address three fundamental questions in the adenovirus physiology. (i) The fiber capsomer being such a complex structure, was it capable of self-assembling, or did it require the presence of a partner protein, acting as a scaffold or chaperon in its trimerization process? (ii) In the latter case, was this partner protein of viral or cellular origin? (iii) Did fiber trimerization preferentially take place in the cytoplasm or within the nucleus, at specific morphogenesis sites related to the nuclear matrix? We performed *in vitro* translation experiments of Ad2 fiber mRNA and analyzed the resulting products with regards to fiber polypeptide synthesis and their assembly into trimers. Four different physicochemical and biological criteria were used to assess the trimeric nature of the fiber synthesized *in vitro* in reticulocyte lysate. (i) Sedimentation coefficient, (ii) electrophoretic mobility of nondenatured samples in SDS-gels, (iii) resistance to proteolysis, and (iv) assembly with the penton base.

EXPERIMENTAL PROCEDURES AND RESULTS

1) Ultracentrifugation Analysis of Ad2 Fiber Synthesized in Vitro; 6 S Trimer Versus 3 S Monomer—Total mRNAs were extracted from KB cells infected with Ad2 at 25 plaque-forming units/cell and harvested at 18 h after infection. Poly(A⁺) mRNAs were isolated by affinity chromatography on an oligo(dT)-cellulose column according to conventional technique (17). The mRNA coding for the fiber polypeptide was purified by two cycles of hybridization on *BclI* fragment D from Ad2 DNA. Rabbit reticulocyte hydrolysate was programmed with fiber mRNA and incubated in the presence of [^{35}S] methionine for various periods of incubation at 30 °C and the resulting *in vitro* translation products fractionated by sedimentation through a preformed sucrose gradient. Each fraction was assayed for the presence of a 62-kDa fiber polypeptide band. After denaturation in SDS-2-mercaptoethanol sample buffer at 100 °C, the gradient fractions were analyzed by standard denaturing SDS-polyacrylamide gel electrophoresis (SDS-PAGE), using a conventional discontinuous buffer system (18). The gels were dried and autoradiographed or

* This work has been supported by grants from the Association pour la Recherche sur le Cancer, the Fondation pour la Recherche Medicale, the Conseil Régional Languedoc-Roussillon, and la Fédération des Groupements des Entreprises Françaises pour la Lutte contre le Cancer. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom all correspondence should be addressed.

[1] The abbreviations used are: Ad, adenovirus; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; NDS-PAGE, nondenaturing SDS-PAGE.

electrically transferred and immunoblotted using anti-fiber serum.

A labeled band of 62 kDa was found in fractions corresponding to proteins sedimenting at 6 and 3 S, the sedimentation coefficient values reported for the native trimeric fiber (3, 12) and for the fiber polypeptide unit (19, 20), respectively (data not shown). This 62-kDa band reacted with fiber antiserum on immunoblots (not shown). The proportion of fiber sedimenting at 6 S was estimated by densitometric scanning of the 62-kDa polypeptide band on autoradiograms and scintillation counting of 62 kDa bands excised from the gels. It increased gradually with the time of *in vitro* translation reaction, and reached a plateau at 30–40% of the total polypeptide synthesized at 2 h. These results suggested that a significant amount of the fiber synthesized *in vitro* was in the trimeric form, as in the naturally occurring fiber capsomer (5).

Since it was important to determine whether *in vitro* assembly was uniquely for adenovirus fiber or whether other structural oligomeric proteins might also assemble after *in vitro* synthesis in our system, we analyzed the *in vitro* translation products of other late mRNAs, *i.e.* the mRNA coding for hexon and the mRNA for penton base. We found that hexon polypeptides synthesized *in vitro* only assembled efficiently in translation reactions containing the scaffolding 100-kDa protein, shifting in sedimentation from 5 (hexon monomer) to 12 S (hexon trimer). On the contrary, penton base monomers (4 S) failed to assemble into pentamers (9 S) under the same conditions, even in the presence of translation products from the other major late mRNAs (not shown). A second electrophoretic criterium for fiber trimerization was developed to overcome the low resolution of 6 and 3 S fiber species in sucrose gradients and to obtain a better estimation of the quantity of fiber trimers than by assaying the 62-kDa fiber polypeptide present in 6 S fractions analyzed by standard denaturing SDS-PAGE.

2) Electrophoretic Mobility of Fiber in Nondenaturing SDS-Polyacrylamide Gel Electrophoresis (NDS-PAGE)—When samples are not denatured by heating in SDS-2-mercaptoethanol-containing buffer prior to the electrophoresis in a standard SDS-polyacrylamide gel (18), the proteins still migrate according to their relative molecular mass, since SDS is present in both gel and buffer, but the quaternary structure of stable protein oligomers is preserved. We termed this type of analysis "nondenaturing SDS-polyacrylamide gel electrophoresis" (NDS-PAGE). In such a system, spontaneously occurring fiber monomers or fiber capsomers denatured with 2% SDS at 100 °C for 1–2 min show a single polypeptide band at 62 kDa, whereas native trimeric fiber capsomers migrate as a 180-kDa protein species (Fig. 1*A*). By combining the resolution of both the ultracentrifugation and the electrophoresis of native proteins in SDS-gels, we were able to more accurately discriminate between assembled (trimeric) and unassembled (monomeric) fibers. *In vitro* synthesized fiber samples were first fractionated through a sucrose gradient and then gradient fractions analyzed by NDS-PAGE. The 6 and 3 S fractions were assayed for the occurrence of 180- and 62-kDa protein species, respectively.

Fig. 1*A* shows that the gradient fractions corresponding to components sedimenting at 6 S contained a major protein band comigrating with the band of control native fiber at 180 kDa. The fractions at 3 S almost exclusively contained monomeric fiber migrating as a 62-kDa species. Both the 180- and 62-kDa bands reacted on blots with a polyclonal antiserum directed toward the N-terminal tridecapeptide of the fiber (Fig. 1*B*). The 180-kDa band strongly reacted with a polyclonal rabbit anti-fiber serum raised against native fiber and adsorbed against heat-SDS-denatured fiber immobilized on an insoluble polyacrylamide support (not shown). In addition, only the 180-kDa band reacted with the monoclonal antibody 2A6–36 specific for fiber trimers (data not shown). The 62-kDa fiber polypeptide molecules sedimenting at 3 S occurred spontaneously and did not result from heat-SDS-denaturation of higher molecular mass oligomeric species. They likely represented an heterologous population of unassembled fiber monomers, *e.g.* nascent fiber polypeptide chains conformationally incompetent for trimerization or fiber monomers at different intermediate stages of folding and maturation (see below).

Scanning of autoradiograms such as in Fig. 1*A* yielded concentrations for fiber trimers ranging between 10 and 15% after 2 h of synthesis *in vitro* and attaining 20% of the total radioactivity recovered in the fiber polypeptide bands after 4 h. These values were somewhat inferior to the estimations made from the radioactivity found in the 6 S fractions of ultracentrifugation gradients or in the 62-kDa band excised from gels after denaturation and standard SDS-PAGE of the same fractions. This was expected, considering the lack of resolution of components sedimenting with close S values. Our two-step analysis, taking advantage of their differences in sedimentation coefficient and electrophoretic mobility in NDS-PAGE, could be considered as the reference technique for assaying fiber monomers and trimers simultaneously in the same samples.

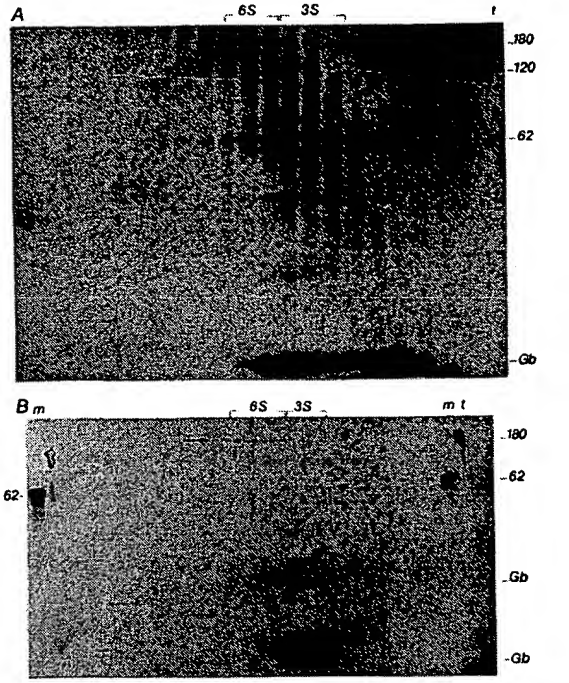

FIG. 1. **Sedimentation analysis of Ad2 fiber synthesized *in vitro*.** *A*, autoradiogram; *B*, immunoblot. 0.1-μg aliquots of fiber mRNA (1–2 μl) were incubated with 30 μl of reticulocyte lysate (Promega), and 3 μl of [^{35}S]methionine (15 μCi/μl; 800 Ci/mmol) for 2 h at 30 °C. Reaction was stopped by cooling in ice and the samples loaded on top of a preformed linear sucrose gradient (5–27%; w:v). After centrifugation for 16 h at 55 krpm and 4 °C in a Kontron TST-60 rotor, 0.2-ml fractions were collected. Native proteins present in each fraction were analyzed by NDS-PAGE. Gel (15% polyacrylamide; acrylamide to bisacrylamide ratio of 100:0.5) was electrophoresed in a standard discontinuous buffer system (18), dried and autoradiographed (*A*) or electrically transferred to nitrocellulose membrane (*B*). The blot was reacted with a polyclonal antiserum raised in rabbit against fiber N-terminal tridecapeptide (6). *Lane m*, control Ad2 fiber monomer (62 kDa) used as marker; *lane t*, native fiber trimer (180 kDa) marker. Fiber trimers sedimenting as 6 S components in sucrose gradient are indicated by *arrows* and fiber subunits (monomers) with an apparent coefficient of 3 S by *arrowheads*. The band at 120 kDa likely represented fiber dimers, sedimenting at an intermediate S value. *In vitro* synthesis of endogenous globin chains (*Gb*), sedimenting at 4 S, resulted in the occurrence of two radioactive protein bands at 16 kDa (monomer) and 32 kDa (dimer) on the autoradiogram (*A*) and two corresponding red-colored bands on the blot (*B*). *Bottom* of the gradient on the *left*.

3) Resistance to Trypsin Digestion—It is known that the fiber capsomers isolated from adenovirus-infected cell lysates are resistant to trypsin digestion, whereas SDS-denatured fiber or fiber polypeptide units are readily hydrolyzed by trypsin (2). *In vitro* assembled fiber trimers were thus assayed by the resistance of the 62-kDa fiber polypeptide band to trypsin digestion. Aliquots from *in vitro* translation reaction were removed at different time intervals and analyzed in standard SDS-PAGE after SDS-2-mercaptoethanol denaturation at 100 °C. Quantification of fiber monomers and trimers was performed by densitometric scanning of the 62-kDa fiber band present in gel autoradiograms of untreated and trypsin-treated samples. As shown in Fig. 2*A*, the band of 62 kDa fiber polypeptide visible in control undigested samples already represented 30% of the maximal synthesis at 20 min of incubation at 30 °C and progressively increased until it reached a plateau at 2 h.

Conversely, no band of trypsin-resistant fiber could be detected in significant amounts before 40 min (Fig. 2*B*), and the radioactivity in the 62-kDa fiber polypeptide band could only be determined after

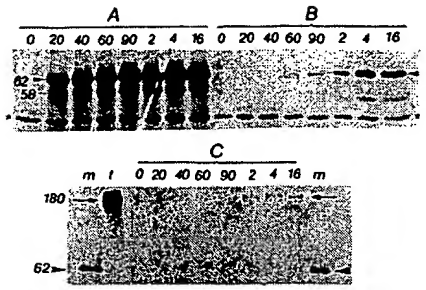

FIG. 2. **Trypsin-resistance assay for Ad2 fiber synthesized *in vitro*.** *A* and *B*, autoradiograms. *C*, immunoblot. Aliquots (3 μl) of fiber mRNA-programmed *in vitro* translation reactions in the presence of [^{35}S]methionine were taken at different time intervals (*lanes*: 20, 40, 60, and 90 min; 2, 4, and 16 h) and further incubated for 1 h at 37 °C with (*B*) or without (*A*) trypsin at a final concentration of 10 μg/μl. After denaturation by heating at 100 °C in SDS-2-mercaptoethanol-containing buffer, the samples were analyzed by standard denaturing SDS-PAGE (18). *Lane 0*, no exogeneous mRNA added. Fiber polypeptide (62 kDa) is indicated by an *arrowhead* and the band of endogeneous protein synthesis activity at 50 kDa by an asterisk. The band at 58 kDa could represent a cleavage product of the fiber polypeptide or a re-initiation at one of the two ATGs coding for the methionine residues at positions 61 or 66 in the fiber sequence. This band disappeared upon trypsin digestion (*B*). In *C*, the same undigested samples as in *A* were kept nondenatured and native proteins analyzed by NDS-PAGE, as in Fig. 1. After electric transfer, the blot was reacted with an anti-N-terminal peptide serum (reacting with both trimers and monomers) and a phosphatase-labeled anti-rabbit IgG conjugate. Fiber trimer (180 kDa) is indicated with an *arrow* and fiber monomer (62 kDa) by an *arrowhead*. *Lane m*, fiber polypeptide marker; *lane t*, native fiber trimer. Note the chase of the signal from the 62- to the 180-kDa band occurring between 90 min and 16 h.

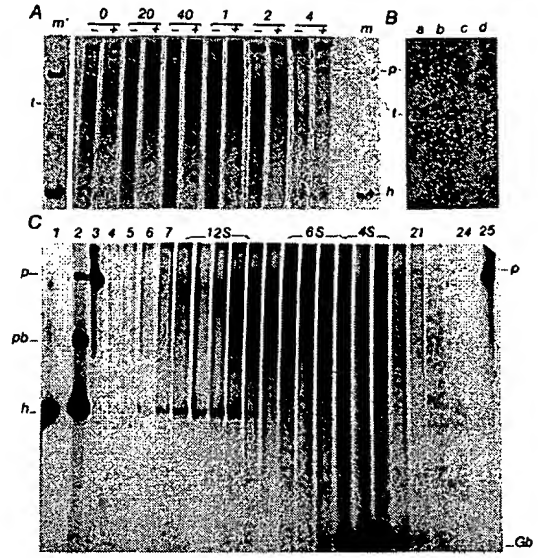

FIG. 3. **Binding of *in vitro* synthesized fiber with penton base.** *A* and *C*, autoradiograms. *B*, immunoblot. In *A*, aliquots (3 μl) of fiber mRNA-programmed *in vitro* translation reaction samples, corresponding to 300 pg of [^{35}S]methionine-labeled fiber polypeptide, were taken at different time intervals (*lanes*: 20 and 40 min; 1, 2, and 4 h) and incubated with (+) or without (−) unlabeled adenovirus penton base for 16 h at 4 °C. The final penton base concentration (1 μg/sample) corresponded to a 3×10^3-fold excess of penton base over the fiber. Proteins were analyzed in their native state in a 10% polyacrylamide gel electrophoresed under nondenaturing conditions, *i.e.* without SDS (7). *Lane m*, [^{14}C]valine-labeled native capsid proteins of Ad2 used as markers: *h*, hexon; *p*, penton (*i.e.* penton base + fiber trimer); *t*, fiber trimer. *Lane m'* is a longer time exposure of *lane m*. Note that a discrete band of labeled component co-migrating with Ad2 penton became discernible over the background in the sample incubated for 1 h with penton base, reaching a maximum after a 4-h incubation. In *B*, 3-μl aliquots of the fiber sample taken at 2 h of *in vitro* translation reaction were incubated with increasing quantities of penton base and analyzed by electrophoresis in a nondenaturing 10% polyacrylamide gel as in *A* and immunoblotted with rabbit serum against native fiber. *Lane a*, no penton base; *lane b*, 1; *lane c*, 2.5; *lane d*, 5 μg, *viz.* 3, 7.5, and 15×10^3-fold excess of penton base, respectively. In *C*, a 30-μl sample of *in vitro* translation reaction, programmed with total Ad2 poly(A$^+$) mRNAs (10 μg) in the presence of [^{35}S]methionine for 2 h at 30 °C, was analyzed by ultracentrifugation in sucrose gradient, as in Fig. 1. Each fraction (50-μl aliquot, 40 μg of protein) was then incubated with unlabeled H2*ts*125 penton base (0.3 μg) for 16 h at 4 °C, and native protein complexes were analyzed by electrophoresis in a 10% polyacrylamide gel, under nondenaturing conditions (without SDS, as in *A*), and autoradiography. *Lanes 1, 2, 3*, and *25*, Ad2 capsid proteins used as markers; *lane 1*, hexon (*h*); *lane 2*, hexon, penton base (*pb*) and penton (*p*); *lanes 3* and *25*, penton. *Lanes 4–24*, nondenatured gradient fractions. Hexon capsomers sediment at 12–13 S, penton capsomer at 9–11 S, fiber capsomer at 6 S, and endogeneous globin (*Gb*) at 4 S. Note the band of labeled penton in *lanes 14–17*, corresponding to fractions where 6 S fiber capsomer sedimented.

prolonged autoradiographic exposure of the films. The trypsin-resistant fiber band was found to represent 1% of the total fiber polypeptides synthesized at 40 min and less than 5% at 60 min (*cf.* Fig. 4). This proportion increased gradually after 90 min from 10% at 2 h to a maximum of 20–30% at 4 h (Fig. 2*B*).

A trypsin-resistant band migrating as a 180-kDa species was also found in the same samples analyzed by NDS-PAGE and immunoblotted with an anti-fiber serum (Fig. 2*C*). This confirmed the trimeric nature of trypsin-resistant fiber. The values for fiber trimers obtained from the trypsin resistance assay were close to the data from our reference combined technique described in Section 2 and depicted in Fig. 1. Due to its simplicity, the trypsin resistance assay was preferentially used for kinetic studies (*cf.* Section 5 and Fig. 4).

4) Fiber Assembled in Vitro Is Competent for Interaction with Penton Base and for Penton Capsomer Formation—In a previous study, we have shown that penton base and fiber can interact and assemble into penton *in vitro* (7, 21), with an association constant of 2×10^7 M^{-1}, in terms of fiber molarity (21). Although the fiber oligomers assembled *in vitro* fulfilled the physicochemical criteria established for trimers, it was of importance to determine whether these trimers were biologically functional, *i.e.* capable of assembling with penton base to form penton. With this aim, aliquots of *in vitro* translation reactions using fiber mRNA were withdrawn at various time intervals and incubated with or without unlabeled penton base. Penton base was isolated from cells infected with fiber-defective H2*ts*125 mutant at 39 °C (21–24). Oligomeric proteins and protein complexes were then analyzed in their native state by polyacrylamide gel electrophoresis, in nondenaturing conditions, *i.e.* without SDS, as previously described (7).

Fig. 3*A* shows an autoradiogram of the gel. A discrete protein band, co-migrating with the band of control penton capsomer, became clearly visible over the background in samples taken after 1 h of cell-free translation reaction and was maximum after 4 h. This protein was identified to penton capsomer since it contained a fiber moiety, as shown by electric transfer of the gel and reacting the blot with anti-fiber antiserum (Fig. 3*B*). Furthermore, a radioactive band of 62 kDa, identified as fiber polypeptide, was found in the same samples (as shown in *A*, *lanes 2* and *4*) immunoprecipitated with anti-penton base serum and analyzed by standard SDS-PAGE and autoradiography (not shown).

Fiber synthesized *in vitro* was therefore competent for penton formation. However, penton base did not seem to serve as scaffold or chaperon protein for fiber trimerization. Only fiber trimers, and not monomers, were found capable of binding with penton base to form penton capsomer. This was shown in the next experiment. *In vitro* translation products obtained with total Ad2 poly(A$^+$) mRNAs in the presence of [^{35}S]methionine were analyzed by ultracentrifugation in sucrose gradient. Each gradient fraction was then incubated with unlabeled H2*ts*125 penton base, and the resulting protein complexes and oligomers were analyzed, along with native Ad2 capsid proteins, by electrophoresis in a nondenaturing non-SDS-containing polyacrylamide gel, and autoradiography.

A band of labeled penton protein (unlabeled in its penton base

domain and labeled in its fiber) was found only in fractions 6 S, where fiber trimers sedimented, and not in fraction 3 S, the position of fiber monomers (Fig. 3*C*). All these results suggested that Ad2 fiber synthesized and assembled *in vitro* was capable of reconstituting penton by specific interaction with its natural penton base partner. Penton assembly did not occur with fiber monomer, but seemed to require pre-existing fiber trimer, as well as *in vivo* synthesized and assembled penton base.

5) Efficiency of Fiber Assembly in Vitro; Kinetics of Synthesis and Trimerization—Fiber trimerization efficiency was estimated by the trypsin resistance assay, as in the experiment presented in Fig. 2. The respective amounts of total fiber polypeptides synthesized and of fiber polypeptides trimerized *in vitro* were quantitatively estimated from gel autoradiogram densitometric scannings of the 62-kDa fiber band present in untreated and trypsin-treated *in vitro* translation samples analyzed in standard SDS-PAGE. The values obtained were plotted *versus* the time of *in vitro* translation reaction. The results of a typical quantitative analysis are shown in Fig. 4.

Fiber polypeptide was detected in quantitatable amounts as early as 10 min after translation reaction at 30 °C, reaching a plateau between 100 and 120 min. The $t_{1/2}$ for monomer synthesis was found to be 40 min. As mentioned above, no fiber trimer was detected before 40 min, and the curve of trimer formation was at a plateau after 4 h. At this time, trimers only represented 25–35% of the total fiber polypeptides synthesized. The $t_{1/2}$ for trimerization of fiber monomers was estimated from the difference between the time corresponding to half-plateau (120 min) and the 40-min lag period, *i.e.* 80 min (1.3 h).

From 100 ng of fiber mRNA translated for 2 h at 30 °C in a 35-μl reaction mixture, the concentration of 62-kDa fiber polypeptide obtained ranged from 50 to 150 pg·μl^{-1}, *viz.* an average value of 3.5 ng/sample. The rate of synthesis of fiber polypeptide chains, as calculated from the slope of the curve in its linear portion between 20 and 60 min, was 18×10^{-18} mol·min^{-1}·μl^{-1}, whereas the rate for trimerization, expressed as the quantity of fiber polypeptide chains acquiring trypsin-resistance between 90 and 240 min, was 3×10^{-18} mol·min^{-1}·μl^{-1}. Thus the rate of the trimerization appears to be six times lower than the rate of fiber polypeptide synthesis.

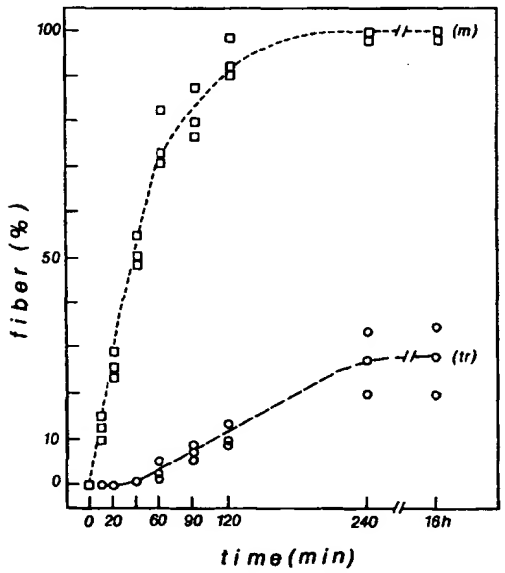

FIG. 4. **Kinetic analysis of *in vitro* fiber synthesis and assembly.** The quantities of total fiber monomers synthesized (*curve m*) and of assembled trimers (*curve tr*) were determined using the trypsin resistance assay (refer to Section 3). Samples of *in vitro* translation reaction were analyzed by standard denaturing SDS-PAGE and the radioactivity in the 62 kDa fiber band estimated before and after trypsin digestion. Radioactive labeling of 62 kDa was quantitated by densitometric scanning of autoradiograms such as those presented in Fig. 2. To prevent misevaluations of counts within overexposed bands, autoradiograms were scanned after different times of exposure, using an automatic scanner (EDC system, Helena Laboratories, Rockville, IL). From the number of methionine residues in the fiber sequence (11) and the specific activity of the methionine used (800 Ci/mmol), the 100% value for the three experiments shown in the graph represented 2.4, 1.6., and 0.8×10^{-15} mol·μl^{-1}, in terms of fiber polypeptides, respectively. The average value of 1.6 was then taken for further calculations of the rate of fiber polypeptide synthesis and assembly.

The apparent 40-min time lag for the occurrence of the first detectable fiber trimers in the curve of trimerization kinetics has several, but not mutually exclusive, interpretations. (i) The fiber polypeptide had to undergo folding to acquire a conformation ready for assembly into a more complex oligomeric structure. (ii) Intermediate steps in assembly control the rate of trimer formation. (iii) The curve of trimer formation shown in Fig. 4 could be considered as being sigmoidal rather than being linear with a lag. In this case, the fiber trimerization would be a cooperative phenomenon; the binding of two monomers would facilitate the binding of a third fiber polypeptide chain to achieve the trimer. This implies that dimers would represent an intermediate step in fiber trimerization.

In this latter hypothesis, some intermediate products with migration intermediate between the 62-kDa monomers and the 180-kDa trimers should be found in the samples of *in vitro* translation reactions. Indeed, a band of 120–130 kDa, a molecular mass fitting with that of a fiber polypeptide dimer and reacting with anti-fiber antibody, was consistently observed on autoradiograms and immunoblots of samples electrophoresed in nondenaturing SDS-gels (Figs. 1*A* and 2*C*). The components corresponding to this band sedimented in sucrose gradient with a coefficient intermediate between the 3 S monomer and the 6 S trimer (Fig. 1*A*). Bacterially expressed Ad2 fiber accumulates as inclusions in *Escherichia coli* cells. After urea solubilization, some intermediate products behaving as dimers have also been detected upon renaturation (25).[2]

Taking these observations into account, we postulated that fiber polypeptide trimerization was controlled by one (or more) intermediate rate-limiting step(s). As previously shown for other oligomeric protein systems (26, 27), the fiber assembly sequence might be written as a succession of unimolecular folding steps and of bimolecular association reactions according to

$$3M' \xrightarrow{k_1} 3M \rightarrow D + M \xrightarrow{k_2} T \qquad (1)$$

where M′ are monomers in a conformational state preceding the folded conformation (M) required for assembly. D are dimers and T are stable trimers in native quaternary structure. Due to the stability of the achieved trimers, this simplified sequence ignored reverse reactions. Other possible intermediate steps, *e.g.* unstable dimers or trimers, were also omitted. Any of the intermediate steps in the sequential first-order reactions of folding and second-order reactions of bimolecular association might be rate-determining in fiber assembly. According to the uni-bimolecular kinetic model, and assuming that one of the reacting components is present in limiting concentration (*e.g.* dimers compared to monomers), the rate of assembly of fiber would depend on subunit concentrations, and the $t_{1/2}$ for fiber assembly would be given by the following equation:

$$1/t_{1/2} = (k_2/\ln 2) \times [M] \qquad (2)$$

in which k_2 is the second-order rate constant (in M^{-1} s^{-1} units), and [M] is the steady-state concentrations of unassembled fiber monomers which are conformationally competent for interacting in the rate-limiting association step that k_2 refers to (27).

The $t_{1/2}$ of trimer formation was 80 min, as deduced from the curves in Fig. 4. Assuming that all polypeptide monomers were competent for assembly, then Equation 2 would give the value of 1×10^5 M^{-1} s^{-1} for the rate constant for Ad2 fiber subunit assembly. If one estimates, however, that the steady-state concentration of competent fiber polypeptide monomers represented one third of the total, then the value of 3×10^5 M^{-1} s^{-1} would be obtained. Reported values for the $t_{1/2}$ of assembly of other oligomeric proteins varied between 0.02 and 400 h and ranged between 1×10^2 and 4×10^5 M^{-1} s^{-1} for the rate constant k_2 (26, 27).

DISCUSSION

The results of the present study suggested that Ad2 fiber polypeptides could self-assemble and form stable trimeric fibers *in vitro* following their synthesis in a cell-free lysate, in the absence of any other adenoviral protein. The possibility exists, however, that some E3 mRNAs (transcriptionally

[2] J. Chroboczek, personal communication.

mapped between 76 and 86 map units on the Ad2 genome) could have been retained during the step of hybridization selection of the fiber message on Ad2 DNA *BclI*-D fragment (8.17–91.3 map units). Translation of such mRNAs would provide our samples with minor but significant amounts of early E3 protein(s) which might positively or negatively interfere with the trimerization process. However, this could be excluded for the following reasons. (i) Early poly(A$^+$) mRNAs from Ad2-infected cells, harvested at 8 h after infection, were selected by hybridization on the same Ad2 DNA *BclI*-D fragment. When these were added to the *in vitro* translation reactions performed with the late fiber mRNA, no difference in the efficiency of fiber assembly was seen. (ii) Ad2 fiber was capable of self-assembling into trimers in insect cells infected by recombinant baculovirus[3] and in human or monkey cells infected with recombinant vaccinia virus.[4] Our data hence suggest that, unlike hexon polypeptides which assemble into trimers only in the presence of functional virus-coded 100-kDa scaffolding proteins (28–30), no adenoviral scaffolding protein is apparently directly involved in the fiber assembly process.

The rate of assembly for fiber polypeptides was relatively slow, compared with the rate of synthesis, and the assembly process relatively inefficient *in vitro*; the number of fiber polypeptides assembled under our conditions never exceeded 25–30% of the total of fiber polypeptide chains synthesized. Undoubtly, however, the fiber trimers obtained *in vitro* were capable of assembling with pre-existing penton base isolated from virus-infected cells, suggesting that the fiber trimers constituted *in vitro* were not functionally different from the fiber capsomers synthesized and assembled *in vivo*.

A significant time lag of 40 min between the synthesis of fiber polypeptides *in vitro* and the occurrence of the first detectable assembled trimers suggested that fiber polypeptide assembly did not occur co-translationally. Thus adenovirus fiber behaves differently from influenza virus hemagglutinin whose subunits have been found to trimerize within minutes of their synthesis *in vivo*, with a $t_{1/4}$ of 5–10 min. Hemagglutinin trimerization probably takes place in the endoplasmic reticulum (31–33) and is apparently independent of glycosylation (33). In the case of adenovirus fiber, intermediate steps, *e.g.* formation of fiber subunit dimers or a proper folding of the nascent fiber polypeptide chain, might be rate-limiting reactions during assembly.

Adenovirus fiber assembly might be assisted by host cell chaperon protein(s) present in the reticulocyte lysate (27, 34–38). Members of the heat shock protein (hsp) family are good candidates for assuring a proper folding or/and assembly of fiber protein (34, 39, 40), and the existence of intracellular interaction between hsp70 and Ad5 fiber has recently been reported (41). Of the two most represented and conserved cytoplasmic hsp families (hsp70 and hsp90), hsp90 is absent from wheat germ cells (42). In order to test whether proteins belonging to the hsp90 family might be involved in fiber trimerization, we translated the fiber mRNA *in vitro* in wheat germ extracts (Promega, Madison, WI) and assayed the samples for the occurrence of anti-fiber antibody-reacting 180-kDa trimer by NDS-PAGE analysis and immunoblotting. No difference could be discerned from the results obtained with reticulocyte lysates (not shown), suggesting that fiber was also capable of assembling in non-mammalian eucaryotic cell-free extracts. It also suggested that members of the hsp90 family were not chaperons for adenovirus fiber trimerization.

The fact that adenovirus fiber trimers could assemble in reticulocyte lysate implied that assembly of fiber need not necessarily take place in the nuclear matrix, where preferred sites for adenovirus particle morphogenesis have been identified (43). Fiber assembly could then occur within the cytoplasm of the infected cell.

Acknowledgments—We gratefully acknowledge the kind help of J. A. Engler and J. S. Hong in providing them with monoclonal 2A6-36 anti-trimer antibody. We also thank M. L. Caillet-Boudin for helpful discussion at the beginning of this work in Lille (INSERM U-233), S. S. Hong for her critical reading of the manuscript, B. Gay for assistance with photographic reproductions, and D. Petite for cell cultures.

[3] A. Novelli and P. A. Boulanger, manuscript in preparation.

[4] J. S. Hong and J. A. Engler, personal communication.

REFERENCES

1. Nermut, M. V. (1984) in *The Adenoviruses* (Ginsberg, H. S., ed) pp. 5–34, Plenum Publishing Corp., New York
2. Pettersson, U. (1984) in *The Adenoviruses* (Ginsberg, H. S., ed) pp. 205–270, Plenum Publishing Corp., New York
3. Lemay, P., and Boulanger, P. (1980) *Ann. Inst. Pasteur Virol.* **131**, 259–275
4. Caillet-Boudin, M. L., Lemay, P., and Boulanger, P. (1991) *J. Mol. Biol.* **217**, 477–486
5. Devaux, C., Adrian, M., Berthet-Colominas, C., Cusack, S., and Jacrot, B. (1990) *J. Mol. Biol.* **215**, 567–588
6. Caillet-Boudin, M. L., Novelli, A., Gesquière, J. C., and Lemay, P. (1988) *Ann. Inst. Pasteur Virol.* **139**, 141–156
7. Devaux, C., Caillet-Boudin, M. L., Jacrot, B., and Boulanger, P. (1987) *Virology* **161**, 121–128
8. Weber, J., Talbot, B. G., and Delorme, L. (1989) *Virology* **168**, 180–182
9. Defer, C., Belin, M. T., Caillet-Boudin, M. L., and Boulanger, P. (1990) *J. Virol.* **64**, 3661–3673
10. Green, N. M., Wrigley, N. G., Russell, W. C., Martin, S. R., and MacLachlan, A. D. (1983) *EMBO J.* **2**, 1357–1365
11. Boulanger, P., and Devaux, C. (1983) *Biochem. Biophys. Res. Commun.* **110**, 913–918
12. Sundquist, B., Pettersson, U., Thelander, L., and Philipson, L. (1973) *Virology* **51**, 252–256
13. Boudin, M. L., Moncany, M., D'Halluin, J. C., and Boulanger, P. (1979) *Virology* **92**, 125–138
14. van Oostrum, J., and Burnett, R. M. (1985) *J. Virol.* **56**, 439–448
15. Ishibashi, M., and Maizel, J. V. (1974) *Virology* **58**, 345–361
16. Caillet-Boudin, M. L., Strecker, G., and Michalski, J. C. (1989) *Eur. J. Biochem.* **184**, 205–211
17. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Laemmli, U. K. (1970) *Nature* **227**, 680–685
19. Velicer, L. F., and Ginsberg, H. S. (1970) *J. Virol.* **5**, 338–352
20. Wilhelm, J. M., and Ginsberg, H. S. (1972) *J. Virol.* **9**, 973–980
21. Boudin, M. L., and Boulanger, P. (1982) *Virology* **116**, 589–604
22. Boudin, M. L., Lemay, P., Rigolet, M., Galibert, F., and Boulanger, P. (1983) *EMBO J.* **2**, 1921–1927
23. D'Halluin, J. C., Milleville, M., Martin, G. R., and Boulanger, P. (1980) *J. Virol.* **33**, 88–99
24. Martin, J. C., Warocquier, R., Cousin, C., D'Halluin, J. C., and Boulanger, P. (1978) *J. Gen. Virol.* **41**, 303–314
25. Chatellard, C., and Chroboczek, J. (1989) *Gene (Amst.)* **81**, 267–274
26. Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol.* **49**, 117–237
27. Rothman, J. E. (1989) *Cell* **59**, 591–601
28. Cepko C. L., and Sharp, P. A. (1982) *Cell* **31**, 407–415
29. Morin, N., and Boulanger, P. (1984) *Virology* **136**, 153–167
30. Morin, N., and Boulanger, P. (1986) *Virology* **152**, 11–31
31. Copeland, C. S., Zimmer, K.-P., Wagner, K. R., Healey, G. A., Mellman, I., and Helenius, A. (1988) *Cell* **53**, 197–209
32. Gething, M. J., McCammon, K., and Sambrook, J. (1986) *Cell* **46**, 939–950
33. Yewdell, J. W., Yellen, A., and Bächi, T. (1988) *Cell* **52**, 843–852
34. Beckman, R. P., Mizzen, L. A., and Welch, W. J. (1990) *Science* **248**, 850–854
35. Ellis, R. J. (1987) *Nature* **328**, 378–379
36. Ellis, R. J., and Hemmingsen, S. M. (1989) *Trends Biochem. Sci.* **14**, 339–342
37. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988) *Nature* **333**, 330–334
38. Zimmermann, R., Sagstetter, M., Lewis, M. J., and Pelham, H. R. B. (1988) *EMBO J.* **7**, 2875–2880
39. Pelham, H. R. B. (1986) *Cell* **46**, 959–961
40. Schlesinger, M. J. (1990) *J. Biol. Chem.* **265**, 12111–12114
41. Macejak, D. G., and Luftig, R. B. (1991) *Virology* **180**, 120–125
42. Dalman, F. C., Bresnick, E. H., Patel, P. D., Perdew, G. H., Watson, S. J., Jr., and Pratt, W. B. (1989) *J. Biol. Chem.* **264**, 19815–19821
43. Weber, J. M., Déry, C. V., Mirza, M. A., and Horvath, J. (1985) *Virology* **140**, 351–359

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 6349–6353, July 1995
Biophysics

Periodicity of polar and nonpolar amino acids is the major determinant of secondary structure in self-assembling oligomeric peptides

HUAYU XIONG*, BRIAN L. BUCKWALTER†, HONG-MING SHIEH†, AND MICHAEL H. HECHT*‡

*Department of Chemistry, Princeton University, Princeton, NJ 08544-1009; and †Agricultural Research Division, American Cyanamid Company, Princeton, NJ 08540

Communicated by Walter Kauzmann, Princeton, NJ, March 16, 1995 (received for review December 2, 1994)

ABSTRACT **The tendency of a polypeptide chain to form α-helical or β-strand secondary structure depends upon local and nonlocal effects. Local effects reflect the intrinsic propensities of the amino acid residues for particular secondary structures, while nonlocal effects reflect the positioning of the individual residues in the context of the entire amino acid sequence. In particular, the periodicity of polar and nonpolar residues specifies whether a given sequence is consistent with amphiphilic α-helices or β-strands. The importance of intrinsic propensities was compared to that of polar/nonpolar periodicity by a direct competition. Synthetic peptides were designed using residues with intrinsic propensities that favored one or the other type of secondary structure. The polar/nonpolar periodicities of the peptides were designed either to be consistent with the secondary structure favored by the intrinsic propensities of the component residues or in other cases to oppose these intrinsic propensities. Characterization of the synthetic peptides demonstrated that in all cases the observed secondary structure correlates with the periodicity of the peptide sequence—even when this secondary structure differs from that predicted from the intrinsic propensities of the component amino acids. The observed secondary structures are concentration dependent, indicating that oligomerization of the amphiphilic peptides is responsible for the observed secondary structures. Thus, for self-assembling oligomeric peptides, the polar/nonpolar periodicity can overwhelm the intrinsic propensities of the amino acid residues and serves as the major determinant of peptide secondary structure.**

The folded structures of proteins are stabilized by a variety of different features, including hydrogen bonding, van der Waals interactions, electrostatic interactions, the hydrophobic effect, and the intrinsic propensities of amino acid side chains for particular secondary structures (1). In recent years, the importance of each of these types of interactions individually has been probed in model peptide systems and in mutagenically altered proteins (2–15).

The importance of one type of interaction *relative* to another type of interaction has received far less attention. This is not surprising since natural proteins typically are stabilized by the concerted action of numerous interrelated features, and it is not possible to isolate any one of these features from all the others. The study of proteins, however, is no longer limited to natural proteins. It is now possible to construct proteins that are designed entirely *de novo* (16–20). With the ability to design proteins from first principles comes the possibility to design structures by focusing on one type of interaction with the hope that optimizing this type will compensate for the mistakes that may result from an incomplete understanding of other features. Thus, the emerging field of *de novo* protein design highlights the importance of probing the *relative* contributions of different types of interactions.

Among natural proteins, two structural features are common to virtually all globular and water-soluble structures: (*i*) an abundance of hydrogen-bonded secondary structure (α-helices and β-strands) and (*ii*) a hydrophobic interior that excludes aqueous solvent.

While hydrogen-bonded secondary structure is abundant in protein structures, the type (α-helix, β-strand, or turn) and the locations of secondary structure differ substantially from one protein to another. This observation led to early proposals that different local sequences have inherently different propensities to form secondary structures and that these local preferences derive from intrinsic propensities of the individual amino acids (21–23). Models in which these intrinsic propensities play a dominant role in determining global structure are appealing because they suggest that structure is dictated by local interactions and thus might be predicted from a thorough understanding of these intrinsic propensities. This hope motivated a great deal of research aimed at determining the intrinsic propensities of each amino acid for the α-helical and β-sheet structures (2–10). Such studies have shown that different side chains indeed show different preferences for local structures. However, it has not been demonstrated that these local preferences actually dictate the global structures of native proteins. Furthermore, theoretical and experimental work has demonstrated that large collections of *de novo* proteins can be constructed using design principles that ignore intrinsic propensities (18, 24).

The second feature common to all water-soluble proteins is the presence of a compact hydrophobic core. The ubiquitous presence of a hydrophobic core in protein structures coupled with thermodynamic studies on nonpolar molecules in water has led to the view that burial of hydrophobic surface area provides the dominant force driving a polypeptide chain to fold into its compact native structure (1, 24–27).

In this paper we ask: What happens when a sequence is designed in such a way that the intrinsic propensities of the amino acids favor one type of secondary structure, but the periodicity of polar and nonpolar amino acids allows burial of hydrophobic side chains only in the alternative secondary structure? How important are intrinsic propensities *relative* to the burial of hydrophobic surface area? To answer this question we have synthesized and determined the secondary structures of a family of synthetic peptides.

MATERIALS AND METHODS

Peptide Synthesis and Purification. Peptides were synthesized on model 9600 peptide synthesizer (Biosearch, Millipore) using fluorenyl-9-methoxycarbonyl (Fmoc) solid-phase chemistry (28, 29). Fmoc amino acids were purchased from either Peninsula Laboratories or Propeptide (Princeton, NJ). PAL resin

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

‡To whom reprint requests should be addressed.

was purchased from Millipore. This resin yields peptides amidated at their C termini. N-terminal acetylation was performed using acetylimidazole in *N,N'*-dimethyl formamide. Cleavage reactions were carried out in a mixture of 95% trifluoroacetic acid/2.5% anisole/2.5% ethanedithiol. Peptides were purified by reverse-phase HPLC using a Dynamax C_{18} column. The correct molecular weights were verified by electrospray mass spectrometry.

Circular Dichroism (CD) Spectroscopy. Lyophilized peptide powders were dissolved in deionized water at initial concentrations ranging between 0.5 and 1.0 mg/ml. Concentrated buffer was added to give a final buffer concentration of 4 mM sodium phosphate (pH 7). Solutions were stirred for 15 min and insoluble material was removed by centrifugation. Peptide concentrations in CD experiments ranged from 10 μM to 200 μM and were determined from the tyrosine absorbance at 280 nm (30). CD spectra were recorded on an Aviv 62DS spectropolarimeter using an averaging time of 1.0 sec and a step size of 0.2 nm. All measurements were carried out at 20°C, using cuvettes with a 1-mm pathlength.

RESULTS

Peptide Design. To bias the intrinsic propensities of each peptide toward a particular secondary structure—either α-helix or β-strand—we chose two groups of amino acids with distinct conformational preferences. Each group would contain one hydrophobic residue, one negatively charged residue, and one positively charged residue.

Group 1 contains amino acids that have an intrinsic propensity for α-helical secondary structure. The hydrophobic amino acid in group 1 was chosen to be leucine. This was a straightforward choice, since leucine is one of the most frequently found residues in the helices of natural proteins (22). Furthermore, of all the hydrophobic amino acids, leucine has the greatest intrinsic α-helical propensity in the model peptides of Chakrabartty *et al.* (3). The negatively charged residue in group 1 was chosen to be glutamic acid. This was also a straightforward choice, since glutamic acid has a high intrinsic helical propensity while aspartic acid has a much lower helical propensity (3). Glutamic acid also has the largest P_α value (1.51) in the Chou–Fasman statistical compilation of helical preferences in natural proteins (22). The positively charged residue was chosen to be lysine, since it has a high helical propensity in model peptides and in natural proteins (3, 22), and, according to the Chou–Fasman analysis, lysine is found in α-helices significantly more often than is arginine (P_α = 1.16 for lysine, P_α = 0.98 for arginine) (22).

Group 2 is composed of residues that have much lower α-helical propensities and tend to prefer β-structure or reverse turns. The hydrophobic amino acid in group 2 was chosen to be isoleucine. This residue has either the highest or second highest β-sheet-forming propensity in two recent studies (9, 10), and, according to the Chou–Fasman analysis, it is the second most commonly observed residue in the β-structure of natural proteins (P_β = 1.60) (22). The only residue having a higher β-sheet-forming propensity in the model system of Minor and Kim (10) is threonine, which is not hydrophobic. The only residue with a higher P_β value in the Chou–Fasman analysis is valine (P_β = 1.70). We chose isoleucine rather than valine because of its greater hydrophobicity (31). The negatively charged residue chosen for group 2 was aspartic acid. This residue is a poor helix former in model peptides and in natural proteins (3, 22) but is among the most commonly occurring residues in β-turns (P_t = 1.46 for aspartic acid) (22). The positively charged residue chosen for group 2 is arginine. This was not a totally satisfying choice since, although the Chou–Fasman analysis finds arginine to be relatively indifferent to helix formation, more recent studies using model peptides rank this residue as a reasonably good helix former (3).

The sequences of the peptides (Fig. 1) were designed to have a repeating pattern of polar and nonpolar amino acids that would match the structural periodicity either of an α-helix (3.6 residues per turn) or of a β-strand (2 residues per turn). Thus, if a given peptide formed the secondary structure matching its sequence periodicity, then the structure would be amphiphilic (32–34) (see Fig. 2). Such structures could bury nonpolar residues by oligomerization of several amphiphilic units. In essence, removal of hydrophobic surface area from contact with solvent water would depend upon the formation of an amphiphilic secondary structure consistent with the sequence periodicity.

Our strategy was to design four peptides: two peptides would be composed of group 1 amino acids (α-intrinsic propensity), and two peptides would be composed of group 2 amino acids (α-avoiding). Within each pair, one peptide would have a sequence periodicity consistent with amphiphilic α-helices (the "A" peptides), while the other would have a sequence periodicity consistent with amphiphilic β-strands (the "B" pep-

Leu, Glu, Lys *(α-HELICAL INTRINSIC PROPENSITIES)*

PEPTIDE 1A *(α-Helical Periodicity)*

```
Tyr-Leu-Glu-Glu-Leu-Leu-Lys-Lys-Leu-Glu-Glu-Leu-Leu-Lys-Lys-Leu
 •   •   -   -   •   •   +   +   •   -   -   •   •   +   +   •
```

PEPTIDE 1B *(β-Strand Periodicity)*

```
Tyr-Lys-Leu-Glu-Leu-Lys-Leu-Glu-Leu
 •   +   •   -   •   +   •   -   •
```

Ile, Asp, Arg *(NON α-HELICAL INTRINSIC PROPENSITIES)*

PEPTIDE 2A *(α-Helical Periodicity)*

```
Tyr-Ile-Asp-Asp-Ile-Ile-Arg-Arg-Ile-Asp-Asp-Ile-Ile-Arg-Arg-Ile
 •   •   -   -   •   •   +   +   •   -   -   •   •   +   +   •
```

PEPTIDE 2B *(β-Strand Periodicity)*

```
Tyr-Arg-Ile-Asp-Ile-Arg-Ile-Asp-Ile
 •   +   •   -   •   +   •   -   •
```

FIG. 1. Sequences of the peptides. Peptides 1A and 1B are composed of residues with α-helical intrinsic propensities, while peptides 2A and 2B are composed of residues with non-α-helical intrinsic propensities. Within each pair, the A peptide has a sequence periodicity of polar (+ and −) and nonpolar (●) residues matching the structural periodicity of an α-helix, whereas the B peptide has a sequence periodicity matching the structural periodicity of a β-strand.

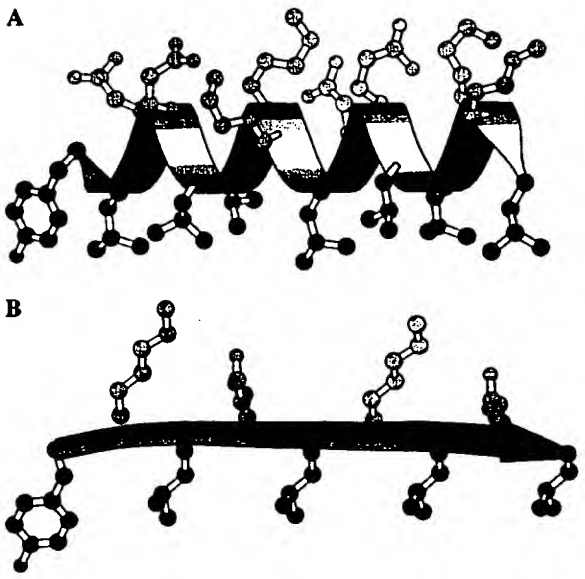

FIG. 2. (*A*) Peptide 1A is shown in α-helical conformation. (*B*) Peptide 1B is shown in β-strand conformation. Polar side chains are shown in gray and nonpolar side chains are shown in black. In each case the sequence periodicity matches the structural periodicity and so the structures are amphiphilic. Diagrams were made using the program MOLSCRIPT (35).

tides). As shown in Fig. 3, the structures formed by each peptide would reflect the relative importance of the intrinsic propensities for a particular secondary structure in competition with the requirement for amphiphilic structure and hydrophobic burial. Peptide "1A" is composed of α-favoring amino acids arranged with the α periodicity, and thus it serves as an α-helical control. Peptide "1B" is also composed of α-favoring amino acids, but they are arranged with the β periodicity; in this peptide the intrinsic propensity and the desire to bury hydrophobic surface area are in direct competition. Peptide "2B" is composed of β-favoring amino acids arranged with the β periodicity and thus serves as a β-strand control. Peptide "2A" is also composed of β-favoring amino acids but arranged with the α periodicity; again, intrinsic propensity competes with hydrophobic burial.

The secondary structures for each peptide as would be predicted either from the intrinsic propensities of the amino

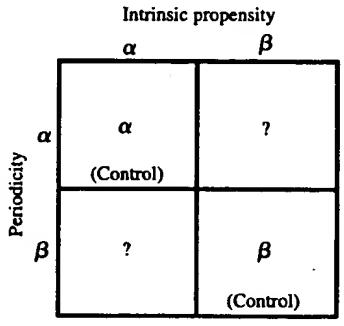

FIG. 3. Schematic diagram showing the competition between intrinsic propensity and polar/nonpolar periodicity. Peptide 1A would be expected to form α-helical structure by both parameters and it serves as the "α control" in the top left. Peptide 2B would be expected to form β structure by both parameters and it serves as the "β control" in the bottom right. For peptides 1B (top right) and 2A (bottom left) there is a competition between the two effects.

acids (22) or from the sequence periodicity are compared in Table 1.

The sequences of the four peptides shown in Fig. 1 are not all the same length but, instead, were designed to have the same number (five) of nonpolar repeats. Ideally, one would also like to compare a set of peptides that were all the same length. However, we found that peptides with alternating polar and nonpolar residues (i.e., the β periodicity) were extremely "sticky;" longer versions of these peptides typically formed gels, suggesting extensive aggregation of β-sheet structure (34, 36, 37) (see *Discussion*).

Each sequence has a tyrosine residue at its N terminus so that peptide concentration could be determined spectrophotometrically (30). Peptides 1A and 2A are acetylated at their N termini and amidated at their C termini. Peptides 1B and 2B are amidated at their C termini. Versions of peptides 1B and 2B that were blocked at both ends were not soluble and so only the singly blocked versions of these peptides were used for further study (see *Discussion*).

Secondary Structures of Peptides. The secondary structures of the designed peptides were probed by CD spectroscopy. This technique is particularly diagnostic for distinguishing between α-helices and β-sheets. The signatures of α-helical structure are two minima at 208 and 222 nm, and a crossover at 200 nm. In contrast, the CD spectrum of β-sheets is dominated by a single minimum at 217 nm and a crossover of ≈205 nm (38, 39). The CD spectra of the four model peptides are shown in Fig. 4. Peptides 1A and 2A are clearly α-helical, while peptides 1B and 2B are clearly β-sheet. Thus, in all cases, the observed secondary structure (Table 1, column 5) correlates with the *periodicity* of the peptide sequence (Table 1, column 4), even when this secondary structure differs from the structure that would have been predicted from the intrinsic propensities of the component amino acids (Table 1, column 3).

Concentration Dependence of Secondary Structure. The results shown in Fig. 4 demonstrate that the periodicity of polar and nonpolar amino acids dominates the decision between α-helices and β-strands. According to the design, these observed secondary structures would be stabilized by the burial of hydrophobic surface area, which accompanies oligomerization of the amphiphilic peptides. Therefore, one would expect the observed secondary structures to be concentration dependent. Fig. 5 shows this is indeed the case. All four peptides show greater secondary structure at high concentrations. Furthermore, the spectra of peptides 1A and 2A show an isosbestic point at 202 nm, indicating a concentration-dependent equilibrium between random coil and α-helix; the spectra of peptides 1B and 2B show an isosbestic point at 205 nm, indicating an equilibrium between random coil and β-sheet.

Peptide 1A shows the least concentration dependence. This is not surprising, since this peptide is composed of strong α-helix formers arranged with an α periodicity, and related peptides are known to form stable four-helix bundles even at relatively low concentrations (40).

DISCUSSION

The results presented above demonstrate that the choice between α-helical and β-sheet secondary structure is controlled by the sequence periodicity of polar and nonpolar amino acids. Thus, although amino acid residues may differ in their intrinsic preferences for one secondary structure versus another (2–10, 21, 22), these preferences can be overwhelmed by the drive to form amphiphilic structures capable of burying hydrophobic surface area. These results corroborate theoretical studies by Dill and coworkers (1, 24, 27), who showed that secondary structures arise from the requirement to bury hydrophobic moieties. Dill's theoretical work and the experiments described here demonstrate that the precise identity of a residue at a particular location in a sequence may be less

Table 1. Secondary structure for each peptide

Peptide	Composition	Intrinsic propensity*	Periodicity	Result
1A	Leu, Glu, Lys	α (P_α = 1.28; P_β = 0.90)	α	α
1B	Leu, Glu, Lys	α (P_α = 1.25; P_β = 0.90)	β	β
2A	Ile, Asp, Arg	β (P_β = 1.14; P_α = 0.97)	α	α
2B	Ile, Asp, Arg	β (P_β = 1.16; P_α = 0.96)	β	β

*Based on the method of Chou and Fasman (22).

important than the simple choice of whether it is polar or nonpolar. This helps to explain why a given fold can be encoded by many different amino acid sequences (41, 42).

While the observed secondary structures of our model peptides correlate with the sequence periodicity of polar and nonpolar amino acids, it must be noted that the structures also correlate with the relative sequence locations of positive and negative charges. Thus, peptides 1A and 2A have charged residues at spacings that would allow formation of multiple salt bridges in α-structure, while peptides 1B and 2B have charged residues at spacings consistent with multiple salt bridges in β-structure. Although formation of salt bridges may contribute to the stabilities of the observed structures, several lines of evidence indicate that salt bridges are not essential:

(*i*) Peptides 1A and 2A are α-helical, not only at pH 7 as described above but also at pH 4.0, where salt bridges would not form (data not shown).

(*ii*) A 13-residue peptide (sequence: Tyr-Arg-Ile-Arg-Ile-Arg-Ile-Arg-Ile-Asp-Ile-Asp-Ile-amide) with the alternating periodicity of polar and nonpolar amino acids forms β-structure, despite having a sequence that is not consistent with the formation of multiple salt bridges (unpublished results).

(*iii*) Previous work with periodic amphiphilic sequences showed that salt bridges are not required for secondary structure formation. For example, early studies demonstrated that poly(Val-Lys) formed β-structure (43). Further studies using either long copolymers (44) or short peptides (33, 45) demonstrated that sequences containing only leucine and lysine residues form β-structures when the leucine and lysine alternate, but form α-helices otherwise. Since those sequences contained no negatively charged side chains, it is apparent that the periodicity of polar and nonpolar residues can dictate secondary structure even in the absence of salt bridges.

While the periodicity of polar and nonpolar residues dominates the decision between α and β secondary structure, it does not ensure that a designed peptide will be water-soluble. Insolubility is a particular problem for sequences with the β periodicity, and amphiphilic β-strands often aggregate into gels or precipitates (34, 36, 37, 46). β-strands are inherently gregarious: The formation of backbone hydrogen bonds and the burial of hydrophobic side chains require a β-strand to interact with neighboring elements of structure. Thus, a given β-strand might form hydrogen bonds to neighboring strands on either side and hydrophobic interactions to a third strand above or below it. (This is quite different from α-helices, which form hydrogen bonds within the same helix.) Such neighborliness can promote nonspecific open-ended aggregation leading to formation of gels or precipitates (37). The tendency toward uncontrolled aggregation can be stemmed (*i*) by maintaining a net charge on the peptide and (*ii*) by choosing relatively short sequences. Thus, peptides 1B and 2B, which are blocked only at their C termini, have a net charge of +1 and form soluble β-structure, but versions of these peptides blocked at both termini are neutral and are not soluble in water. Likewise, while the 9-mers 1B and 2B formed soluble β-structure, longer variants of these peptides typically formed gels or precipitates (unpublished results).

Our studies with synthetic peptides can be considered in terms of models proposed to describe the folding pathways of native proteins (47). Among these models are the framework model and the hydrophobic collapse model. According to the framework model, the secondary structure of a protein forms before the tertiary structure is locked into place. According to the hydrophobic collapse model, nonpolar residues are sequestered away from aqueous solvent early in folding to yield a "molten globule" state; this state then continues on to the native structure by accumulating the precise tertiary interactions of the native fold (47). Our results suggest that these two models should not be considered as competing alternatives but rather as simultaneous and interdependent occurrences. Early events in folding would resemble oligomerization of our amphiphilic peptides. Thus, the secondary structures of the framework model would be stabilized early in folding because of the hydrophobic collapse of nonpolar faces. If protein folding indeed follows a path of hydrophobic collapse with

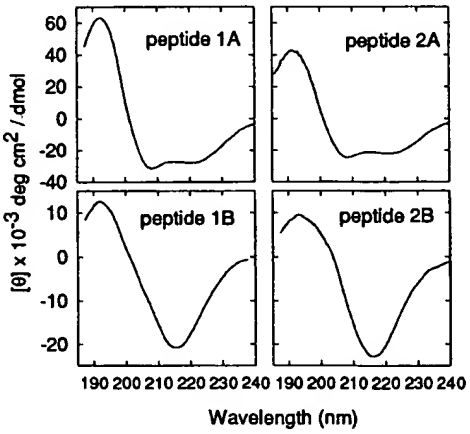

FIG. 4. CD spectra (and concentrations) for peptides 1A (66 μM), 2A (98 μM), 1B (174 μM), and 2B (90 μM). The layout is the same pattern as in Fig. 3.

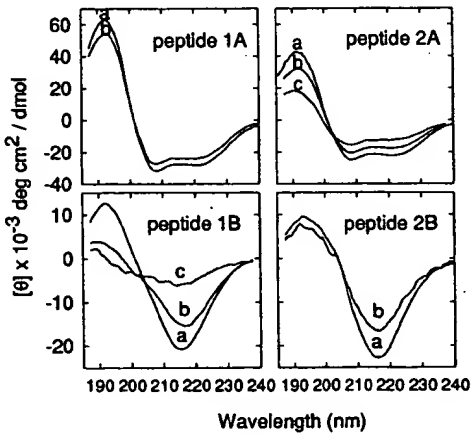

FIG. 5. Concentration-dependent CD spectra for model peptides. Concentrations are as follows: peptide 1A, 13.2 μM and 65.8 μM; peptide 2A, 24 μM, 49 μM, and 98 μM; peptide 1B, 44 μM, 87 μM, and 174 μM; peptide 2B, 31 μM and 90 μM.

concomitant formation of secondary structure, proteins could find their native structures without extensive conformational searching (27). Indeed, the early stages in protein folding may be approximated more closely by these self-assembling peptides than by the monomeric peptides typically used to study intrinsic propensities of local sequences.

Our findings also have implications for *de novo* protein design. The design of novel proteins typically requires a balance between designing *for* the desired structure and *against* all possible alternative structures. To accomplish this task, the would-be designer must consider the various features known to be important in stabilizing proteins. We have shown previously that four-helix bundles can be designed by using a "binary code" with the appropriate sequence periodicity of polar and nonpolar residues (18). Intrinsic propensities were not included as part of that design strategy, and so it was difficult to gauge their importance. Because that collection of sequences was generated by semi-random combinatorial methods, no individual sequence was suitable for explicitly comparing the importance of intrinsic propensities relative to sequence periodicity. The current experiments use defined peptide sequences to demonstrate explicitly that when the periodicity of an amino acid sequence matches the repeat pattern of a secondary structure, the sequence will form that secondary structure—regardless of the intrinsic propensities of the component amino acids. Thus, the periodicity of polar and nonpolar amino acids appears to be a major determinant for *de novo* protein design.

We thank Roy Bass, Billy Bullock, Patrick Mowery, and Judith Swan for advice and assistance. This work was supported by the American Cyanamid/Princeton University Grantlet Program.

1. Dill, K. A. (1990) *Biochemistry* **29,** 7133–7155.
2. Wójcik, J., Altman, K. H. & Scheraga, H. A. (1990) *Biopolymers* **30,** 121–134.
3. Chakrabartty, A., Kortemme, T. & Baldwin, R. L. (1994) *Protein Sci.* **3,** 843–852.
4. Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. (1990) *Science* **250,** 669–673.
5. Merutka, G., Lipton, W., Shalongo, W., Park, S. H. & Stellwagen, E. (1990) *Biochemistry* **29,** 7511–7515.
6. Zhou, N. E., Kay, C. M., Sykes, B. D. & Hodges, R. S. (1993) *Biochemistry* **32,** 6190–6197.
7. O'Neil, K. T. & DeGrado, W. F. (1990) *Science* **250,** 646–651.
8. Kim, C. A. & Berg, J. M. (1993) *Nature (London)* **362,** 267–270.
9. Smith, C. K., Withka, J. M. & Regan, L. (1994) *Biochemistry* **33,** 5510–5517.
10. Minor, D. L. & Kim, P. S. (1994) *Nature (London)* **367,** 660–663.
11. Blaber, M., Zhang, X.-J. & Matthews, B. W. (1993) *Science* **260,** 1637–1640.
12. Horovitz, A., Matthews, J. M. & Fersht, A. R. (1992) *J. Mol. Biol.* **227,** 560–568.
13. Alber, T., Sun, D. P., Wilson, K., Wozniak, J. A., Cook, S. P. & Matthews, B. W. (1987) *Nature (London)* **330,** 41–46.
14. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992) *Science* **255,** 178–183.
15. Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990) *Biochemistry* **29,** 2403–2408.
16. Regan, L. & Degrado, W. F. (1989) *Science* **241,** 976–978.
17. Hecht, M. H., Richardson, J. S., Richardson, D. C. & Ogden, R. C. (1990) *Science* **249,** 884–889.
18. Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M. & Hecht, M. H. (1993) *Science* **262,** 1680–1685.
19. Mutter, M. (1988) *Trends Biochem. Sci.* **13,** 260–265.
20. Ghadiri, M. R., Soares, C. & Choi, C. (1992) *J. Am. Chem. Soc.* **114,** 4000–4002.
21. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13,** 211–222.
22. Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276.
23. Levitt, M. (1978) *Biochemistry* **17,** 4277–4285.
24. Yue, K. & Dill, K. A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4163–4167.
25. Kauzman, W. (1959) *Adv. Protein Chem.* **14,** 1–63.
26. Lau, K. F. & Dill, K. (1989) *Macromolecules* **22,** 3986–3997.
27. Dill, K. A., Fiebig, K. M. & Chan, H. S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1942–1946.
28. Atherton, E. & Sheppard, R. C. (1989) *Solid Phase Peptide Synthesis: A Practical Approach* (IRL, Oxford).
29. Fields, G. B. & Noble, R. L. (1990) *Int. J. Pept. Protein Res.* **35,** 161–214.
30. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954.
31. Fauchere, J. L. & Pliska, V. (1983) *Eur. J. Med. Chem.* **18,** 369–375.
32. Brack, A. & Spach, G. (1981) *J. Am. Chem. Soc.* **103,** 6319–6323.
33. DeGrado, W. F. & Lear, J. D. (1985) *J. Am. Chem. Soc.* **107,** 7684–7689.
34. Osterman, D. & Kaiser, E. T. (1985) *J. Cell. Biochem.* **29,** 57–72.
35. Kraulis, P. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
36. Zhang, S., Holmes, T., Lockshin, C. & Rich, A. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 3334–3338.
37. Hecht, M. H. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8729–8730.
38. Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4116.
39. Johnson, W. C. J. (1990) *Proteins Struct. Funct. Genet.* **7,** 205–214.
40. Ho, S. P. & Degrado, W. F. (1987) *J. Am. Chem. Soc.* **109,** 6751–6758.
41. Bajaj, M. & Blundell, T. (1984) *Annu. Rev. Biophys. Bioeng.* **13,** 453–492.
42. Argos, P. (1987) *J. Mol. Biol.* **197,** 331–348.
43. Brack, A. & Orgel, L. E. (1975) *Nature (London)* **256,** 383–387.
44. Brack, A. & Spach, G. (1981) *J. Am. Chem. Soc.* **103,** 6319–6323.
45. Johnsson, K., Allemann, R. K., Widmer, H. & Benner, S. A. (1993) *Nature (London)* **365,** 530–532.
46. Richardson, J. S. & Richardson, D. C. (1987) *Protein Engineering,* eds. Oxander, D. L. & Fox, C. F. (Liss, New York), pp. 149–163.
47. Kim, P. S. & Baldwin, R. L. (1990) *Annu. Rev. Biochem.* **51,** 631–660.

0022-1767/92/14812-4012$02.00/0
THE JOURNAL OF IMMUNOLOGY

Vol. 148, 4012–4020, No. 12, June 15, 1992
Printed in U.S.A.

DESIGN AND SYNTHESIS OF A SELF-ASSEMBLING PEPTIDE DERIVED FROM THE ENVELOPE PROTEINS OF HIV TYPE 1

An Approach to Heterovalent Immunogens

SRIKANTH P. TRIPATHY,* ANIL KUMAR,† VENKATASAMY MANIVEL,† SUBRAT K. PANDA,‡ AND KANURY V. S. RAO[1†]

*From the *AIDS Unit, National Institute of Virology, Pune, India, †Virology Group, International Centre for Genetic Engineering and Biotechnology, NII Campus, Shaheed Jeet Singh Marg, New Delhi-110 067, India, and ‡Department of Pathology, All India Institute of Medical Sciences, New Delhi-110 029, India*

A chimeric peptide that included sequences from gp120 and gp41 of HIV type 1 was synthesized. Cleavage from solid support yielded a composite of self-oligomerized products with molecular masses ranging from 5 to about 9 kDa. The oligomer but not its reduced, monomeric form was recognized by human anti-HIV sera and at least one of the two lysines in the sequence was involved in antibody binding. The oligomeric peptide was immunogenic, yielding a conformation-specific antibody response. Co-oligomerization of a hepatitis B surface Ag-derived peptide and the HIV type 1-derived peptide yielded a bivalent product in which conformational integrity of the individual components was maintained. Immunization with this hybrid peptide resulted in conformation-specific antibodies to both epitopes in all four murine strains tested. Lymphocyte proliferation assays revealed that the T epitopes resident in both peptide sequences remained active in the hybrid peptide. These results demonstrate the potential of this approach in generating multi- and heterovalent immunogens which may eventually find application as vaccines.

The pioneering observations that synthetic peptides derived from antigenic proteins of a variety of pathogens can elicit a cross-reactive humoral response has led to the distinct possibility of eventually developing peptide vaccines (see Ref. 1 for review). Although much effort has been spent, three major obstacles have severely impeded the progress in this area. The first relates to the Ir gene-linked restriction of immune response (2–4). Given that a small synthetic peptide can contain only a limited set of T cell determinants vis-à-vis the native protein, the problem of genetic restriction becomes particularly daunting in the context of an outbred target population. However, the recent discovery of promiscuous T cell epitopes may offer a way around this problem (5–7).

The second obstacle arises from the poor immunogenicity of a synthetic peptide in comparison with the native protein from which the synthetic peptide was derived (8–10). Though it is true that synthetic peptides are notoriously poor immunogens, it is now becoming apparent that only a fraction of the polyclonal antibody response elicited by a viral Ag is involved in neutralization of infection. Traditionally such epitopes that elicit a neutralizing response have been termed as "neutralization epitopes" (11). In recent experiments it was observed that though antibody titers ranging between $1:10^5$ and $1:10^6$ could be observed by immunizing chimpanzees with gp120, the envelope protein of HIV-1,[2] neutralization titers were 100- to 1000-fold lower (12). Further the neutralization titers correlated well with titers against synthetic peptide representing the principal neutralizing epitope on gp120 (12). The implication of these findings then is that, provided it is a fair representative of the neutralization epitope (or epitopes), it may not be necessary for a synthetic peptide to be as immunogenic as the native protein to elicit a comparable level of protective response. On the contrary, as we have recently shown for the HBsAg system (13), because a peptide mimetic represents a "naked" form of the native epitope it is sometimes possible to obtain a better quality of epitope-specific antibodies than that obtained with the native protein where the relevant epitope may be masked by either glycosylation or protein folding (14, 15).

The third problem often encountered arises from the fact that epitopes expressed by protein Ag frequently result from conformationally restrained stretches of amino acid sequence and can be either contiguous or discontiguous (11, 16–18). This has posed severe restrictions on the ability to generate accurate peptide mimetics of such epitopes, though limited success has been achieved in some systems (19). The flexibility inherent in a short amino acid sequence permits synthetic peptides to occupy multiple conformational states of which only a fraction may be reminiscent of the "native" state of cognate sequence in the parent protein (20). Interestingly in some viral Ag it has been found that the epitopes relevant for protection occur in the loop-regions of the protein where the loop is stabilized by a disulfide bridge (21). In the case of the HBsAg two adjacent loops between residues 124 to 137 and 139 to 147 have been identified

Received for publication December 9, 1991.
Accepted for publication March 20, 1992.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Address correspondence and reprint requests to Dr. Kanury V. S. Rao, International Centre for Genetic Engineering and Biotechnology, NII Campus, Shaheed Jeet Singh Marg, New Delhi-110 067, India.

[2] Abbreviations used in this paper: HIV-1, HIV type 1; HBsAg, hepatitis B surface Ag; HBV, hepatitis B virus; HRPO, horseradish peroxidase; MBHA, 4-methylbenzhydrlamine resin; RT, room temperature.

to each represent a dominant, conformational group specific epitope (21–25). We recently demonstrated that a synthetic peptide spanning the sequences of both loops (residues 124–147) is capable of spontaneous self-assembly via disulfide mediated oligomerization to regenerate a conformational group specific determinant of HBsAg (13). Expression of HBsAg related epitope was found to be stringently dependent upon the integrity of the oligomer and even partial reduction resulted in near complete abolition of antigenicity (13).

A notable feature of the 124–147 sequence of HBsAg is its cysteines. There are in all five cysteine residues, one each at the amino and carboxy termini and three consecutive cysteines toward the middle (13). We believe that it is the odd number of cysteines which facilitates the intermolecular oligomerization, a property which is also common to the parent HBsAg protein (26). An analysis for putative structural elements in the sequence indicated the presence of two overlapping tetrapeptide sequences around the three cysteine central portion which have high probability of existing in a β-turn. Three additional tetra-peptide stretches between 129–132, 142–145 and 144–147 also exhibit this property. We speculated that these turn sequences may play a critical role in the folding of peptide during oligomerization in a native protein like manner (13).

We were interested in determining whether these observations could be extended to other viral Ag systems. In this report we demonstrate that a designed synthetic peptide that includes sequences from gp120 and gp41, the envelope proteins of HIV-1, also spontaneously self-oligomerizes to regenerate a dominant, conformational epitope resident on the native Ag. The approach was further exploited by a co-oligomerizing the HBsAg and HIV-1 derived peptides to obtain a single molecule that expresses conformational epitopes derived from both pathogens. Immunologic analyses revealed that the component epitopes remained codominant both at the B and Th cell level thus providing a powerful and versatile new strategy for the eventual development of peptide based multi and heterovalent vaccines.

MATERIALS AND METHODS

Materials. Human anti HIV-1 sera were obtained from individuals infected with HIV-1 who were positive for antibodies both by ELISA and Western blot criteria with commercially available diagnostic kits. Human anti-HBsAg sera were from individuals who had received three doses of either plasma-purified (H-B-Vax, Merck Sharp & Dohme, West Point, PA) or a rDNA-derived (Engerix B, Smith Kline Biologicals, Rixensert, Belgium) hepatitis B vaccine. All of these sera tested positive for anti-HBsAg with commercially available kits and were also positive for anti-OS antibodies in an ELISA assay as previously described (13).

HRPO-conjugated antibodies to mouse and rabbit IgG were purchased from Sigma Chemical Company (St. Louis, MO, USA). HRPO-labeled goat anti-Human IgG was obtained from the reagent reference center of the National Institute of Immunology (New Delhi, India).

Animals. Mice and rabbits (NZW) were obtained from the breeding facility at the National Institute of Immunology, New Delhi. All mice used for the study were of either sex and between 6 and 8 wk of age.

Peptide synthesis and workup. Peptide HV (see Fig. 1) was synthesised by the solid-phase procedure on MBHA by using an Applied Biosystems model 430A automated peptide synthesizer. Cleavage of peptide from resin and subsequent workup was as previously described (13).

The monomer peptide derivative (peptide MHV) was obtained by reduction with a 10-fold molar excess of dithiothreitol (RT, 3 h) followed by treatment with a 24-fold excess of iodoacetamide (RT, 30 min). The crude product was purified by reverse phase HPLC on a C-18 column (μ Bondapak, 7.8 mm × 30 cm) by using an aqueous gradient of 0 to 60% acetonitrile in 0.1% trifluoroacetic acid over 40 min.

Preparation of peptide HV (K-Me). Four milligrams of peptide HV were dissolved in 2 ml of PBS (pH 7.4) and 20 μl of 30% formaldehyde added. After incubation at 37°C for 15 min, 106 mg of sodium cyanoborohydride were added to the solution and incubated with mixing at 37°C for an additional 30 min. The solution was then dialyzed exhaustively (4 × 2 liters) against deionized water and lyophilized to yield product.

Preparation of composite peptide OS-HV. Aliquots of 250 mg of protected peptides OS on MBHA resin (13) and 250 mg of protected peptide HV-MBHA resin (substitution of each 0.17 mmol/g) were thoroughly mixed and cleaved with hydrogen fluoride in the absence of any reducing thiols (done by Multiple Peptide System, San Diego, CA). The crude peptide was analyzed and purified by cation exchange HPLC on a protein-PAK SP-5PW column (7.5 mm × 7.5 cm) by using a nonlinear sodium chloride gradient in 10 mM sodium phosphate (pH 6.0).

ELISA. Wells were coated with 1 μg of Ag in 100 μl of PBS (pH 7.4) at 37°C for 3 h. Subsequently they were blocked with 200 μl of 5% of fat-free dry milk powder at 37°C for 2 h and washed thoroughly. Then 100 μl of indicated dilution of antibody were added and incubated overnight at 4°C. After washing, wells were incubated with 100 μl of HRPO-labeled secondary antibody at 37°C for 1.5 h. Chromogen used was *o*-phenylenediamine and absorbance was measured at 490 nm.

A sandwich ELISA assay for peptide OS-HV. A 1/500 dilution of mouse mAb specific for the "a" determinant of HBsAg (A 2-73, a kind gift from Dr. J. Pillot, Pasteur Institute, Paris, France) were coated in 100 μl onto wells of an ELISA plate at 37°C for 3.5 h. Wells were blocked as described above. To each well was added 100 μl of dilution buffer (PBS with 0.1% Tween 20) containing 2 μg of peptide and incubated overnight at RT. Wells were washed and to each well was added 100 μl of a 1/20 dilution of human serum that was strongly positive for anti HIV-1 antibodies. This was incubated at 37°C for 1.5 h. After washing, wells were then incubated in 100 μl of a 1/1000 dilution of HRPO-labeled goat anti-human IgG at 37°C for 1 h. Plates were subsequently developed using *o*-phenylenediamine as chromogen and absorbance measured at 490 nm.

Immunization of peptide HV in rabbits. Two NZW rabbits (approximately 1 kg in weight) were immunized i.m. with 200 μg each of peptide in CFA (total volume = 600 μl per rabbit). Six weeks later (day 42), both rabbits were boosted with an identical dose of peptide in IFA. Blood was collected on day 49 from the retro-orbital plexus. Assays of individual sera indicated that, although both the rabbits responded to peptide immunization, anti-HV titer obtained with one of the rabbits was much higher (1:16,000) relative to the other (1:4,000). Serum obtained from the former rabbit was used in all the experiments described here.

Immunization of mice with peptide OS-HV. Groups of four mice from each strain were immunized i.p. with 20 μg of peptide in CFA (total volume 50 μl/mouse). These mice were boosted 3 wk later (day 21) with an identical dose of same peptide in IFA. Mice were bled from the retro-orbital plexus on day 28 and sera from individual mice within a group pooled.

Lymphocyte proliferation assays. Groups of three mice were primed with 20 μg of peptide in CFA (total volume 100 μl). Immunization was intradermal and performed at the base of tail. Seven days later, the mice were sacrificed and inguinal lymph nodes were removed, teased, and washed twice in RPMI 1640. Lymphocytes were resuspended in 1 ml of RPMI 1640 containing 10% FCS and aliquots removed to determine cell counts. The viability of the cells was greater than 95% in all the cases. Cells were appropriately diluted in 100 μl of medium and plated in to wells of 96-well tissue culture plates (Costar, Cambridge, MA) at 5×10^5 cells/well. Aliquots of 100 μl of medium containing 10 μg of appropriate challenge peptide were added in quintuplet and the cells were allowed to grow for 72 h. At the end of this period, each well was pulsed with 1 μCi of [^{3}H] thymidine (Amersham, Buckinghamshire, UK) in 20 μl and the cells incubated for a further 18 h. Cells were then harvested and incorporated radioactivity determined by scintillation counting in a Beckmann LS 3801 counter. Stimulation index was determined as a ratio of the mean of counts per minute obtained in Ag-stimulated wells over that obtained for control wells where no peptide was added.

A 10-μg dose of challenge Ag was selected on the basis of preliminary dose-response standardization experiments with peptide doses of 2, 10, and 50 μg where an optimal response was obtained with 10 μg of peptide in all cases.

RESULTS

Design and synthesis of a self-oligomerizing peptide derived from the envelope proteins of HIV-1. Previous studies with synthetic peptides have identified a domi-

nant epitope on gp41 between residues 598 and 609 (27). Interestingly, this epitope appears to be of the "loop"-type because a disulfide bond between cysteines at position 603 and 609 is mandatory for optimal antigenicity (27).

To extend our previous observations with an HBsAg-derived peptide, we designed and synthesized a chimeric peptide that included sequences from both gp120 and gp41 as shown in Figure 1. The sequences between residues 1 to 14 of this peptide was derived from amino acids 235 to 248 of gp 120 from the BH10 isolate of HIV-1. This sequence is relatively conserved among various isolates and contains a predicted B cell epitope (28). The sequence between residues 16 to 22 of the peptide in Figure 1 represents the gp41 sequence 603–609 previously characterized as a dominant loop-type epitope (27). An additional cysteine was included at position 15 to obtain a cyteine residue distribution comparable to that present in the 124–147 sequence of HBsAg (13).

The resulting peptide (peptide HV) has two overlapping tetrapeptide sequences in the middle (10–13 and 11–14) that have a high probability of occurring in a β-turn ($p_t = 0.6 \times 10^{-4}$ and 1.2×10^{-4}) as per the Chou-Fasman predictive algorithm (29).

In addition there are two other tetrapeptide sequences at 1–4 and 16–19 that can also form β-turns ($p_t = 3.8 \times 10^{-4}$ and 2.8×10^{-4}). Thus with respect to both distribution of cysteines and putative β-turn elements peptide HV emulates the HBsAg 124–147 sequence (13).

Peptide HV was obtained by the stepwise solid-phase synthetic procedure followed by cleavage and workup as described in *Materials and Methods*. Identity of the peptide was confirmed by amino acid analysis where the expected composition was obtained (data not shown). An Ellman analysis for the presence of free sulfhydryl groups (30), revealed that at least 70% of the cysteines were engaged in the disulfide bonds. To ascertain that the disulfide bonds included intermolecular cross-linking (13), peptide HV was resolved by gel electrophoresis under nonreducing conditions (Fig. 2). Peptide HV gave a broad streak that extended from a molecular mass of about 5 kDa to a molecular mass of about 9 kDa with the majority being localized between a range of 6 to 8 kDa. From the theoretical molecular mass of the monomer (2.3 kDa), we infer that peptide HV does indeed represent a heterogenous population of multimeric complexes of which the majority probably represents the trimeric species.

Peptide HV represents a dominant epitope on HIV-1 that is conformation dependent. To ascertain its antigenic properties peptide HV was screened against a total of 99 human serum samples that were pretested to be positive for anti-HIV-1 antibodies with commercially available kits (see *Materials and Methods*). Of these, 93 were also positive for anti-peptide HV antibodies, whereas of the 67 normal human sera used only two showed cross-reactivity with peptide HV (Table I). These results indicate that recognition of peptide HV is specific for anti-HIV-1 sera and that this peptide represents a dominant epitope on the envelope of HIV-1 when presented to the human immune system.

We were next interested in determining whether the

Figure 1. Amino acid sequence of peptide HV. See text for details. The one-letter code for individual amino acid residues has been used.

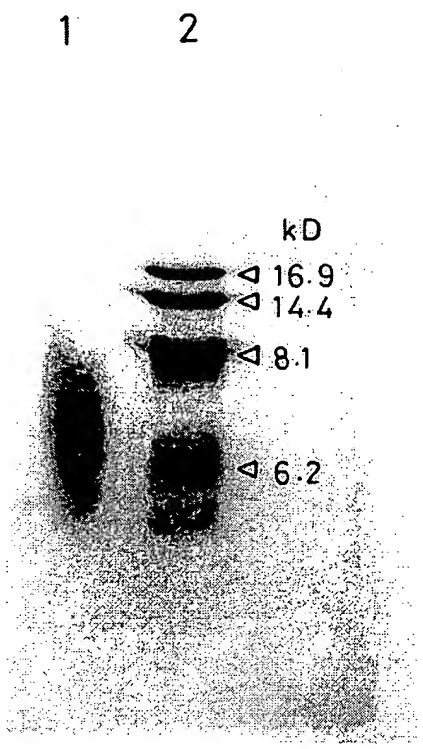

Figure 2. PAGE of peptide HV under nonreducing conditions. Fifteen micrograms of peptide HV in 10 μl were dissolved in an equal volume of 2-ME-deficient sample buffer (8% SDS, 0.25 M Tris, 40% (v/v) glycerol, and 0.005% bromphenol blue) and resolved on an 18% polyacrylamide gel. Visualization was done by staining with Coomassie brilliant blue R 250. Lanes: *1*, Peptide HV; *2*, molecular mass markers (in kDa). Though peptide MHV was also run in parallel, it was not visible by the above mentioned staining protocol.

TABLE I

Peptide HV represents one or more commonly recognized determinant on HIV-1 by the human immune system[a]

Anti-HIV-1 Antibody Status	No. of Samples	Reactivity against Peptide HV	
		Positive	Negative
Positive	99	93	6
Negative	67	2	65

[a] Wells coated with peptide HV were incubated with 100 μl of a 1/10 dilution of either human anti-HIV-1 positive or human anti-HIV-1 negative sera (see *Materials and Methods*). In addition a 1/10 dilution of positive, cut-off and negative control sera from the Wellcozyme HIV recombinant kit (Wellcome Diagnostics, Dartford, England) were also included in duplicate. All test samples that gave an absorbance value higher than that obtained with the cut-off control serum were scored as positive while those with lower absorbance values were scored negative. The positivity or negativity of a given sample was verified by a repeat experiment.

HIV-related epitope expressed by peptide HV was conformation dependent. For this we generated a linear analog of peptide HV by disulfide reduction and subsequent carboxymethylation with iodacetamide (see *Materials and Methods*). In addition to this, to obtain some information on the localization of the epitope, another derivative of peptide was prepared where the lysine residues at positions 4 and 19 were modified by methylation of the side chain amino groups (see *Materials and Methods*). All of these were then screened for reactivity with a representative human anti-HIV-1-positive serum by ELISA and the results are shown in Figure 3.

In comparison with peptide HV its linear analog showed virtually no reactivity indicating that integrity of disulfide bonds is essential for optimal expression of HIV-related antigenicity. Interestingly, a near complete loss in anti-

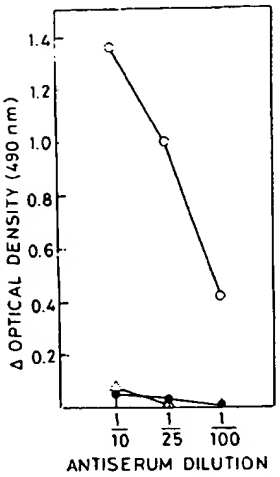

Figure 3. Peptide HV represents a conformational epitope of which a lysine residue represents an important constituent. A representative human HIV-1 antiserum was titrated against a coat of either peptide HV (○), peptide MHV (●), or peptide HV[K-Me] (Δ) at the indicated dilutions in an ELISA protocol (see *Materials and Methods*). Data are presented as the absorbance obtained at a given dilution of test serum minus the absorbance obtained with a negative control serum at the same dilutions and are the mean of duplicate determinations.

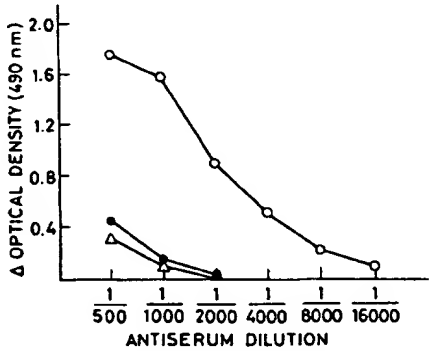

Figure 4. Peptide HV is immunogenic and elicits a predominant antibody response against a conformational epitope that includes a lysine residue. Polyclonal rabbit anti-HV serum was titrated at the indicated dilutions against either peptide HV (○), peptide MHV (●) or peptide HV[K-Me] (Δ) by ELISA. Values presented are the mean absorbance obtained for a given dilution after subtracting the absorbance obtained with preimmune serum at the same dilution and against the same coat Ag. Figure is a representative of three separate experiments.

genicity was also displayed by the peptide derivative in which the lysine side chains had been chemically modified (peptide MHV).

From these results, we infer that the major HIV-related eptiope expressed by peptide HV is dependent on a conformation that is induced by disulfide bond formation and that the epitope includes at least one of the two lysine residues in the sequence. Preliminary results suggest that a significant proportion of the HIV-1-related antigenicity may be due to the sequence derived from gp-41 as a polyclonal antiserum preparation against gp120 (a kind gift from Dr. R. H. Purcell, National Institutes of Health, Bethesda, MD) showed poor cross-reactivity with peptide HV (data not shown).

Peptide HV is immunogenic and elicits an anti-peptide response that is conformation specific. Immunization of two rabbits with peptide HV yielded an anti-peptide response in both rabbits indicating that the peptide is immunogenic in the absence of a carrier protein. This rabbit polyclonal antiserum preparation reacted weakly with the linear analog of peptide HV (Fig. 4) indicating that a majority of antibodies elicited by peptide HV are conformation specific. Interestingly the rabbit anti-peptide serum also reacted poorly with the derivative of peptide HV where the lysine side chains were chemically modified (Fig. 4). Thus peptide HV elicits a predominant antibody response against a conformational epitope that is either the same as or overlaps with the one that is recognized on the corresponding sequence of HIV-1 in humans.

Construction and characterization of a hybrid molecule that includes conformational epitopes derived from HBV and HIV-1. We have so far demonstrated that a designed synthetic peptide derived from the envelope sequence of HIV-1 self-assembles in a disulfide dependent manner to regenerate a dominant conformational epitope present on the native Ag. Furthermore, antibodies elicited in response to peptide HV are predominantly specific for the disulfide bonded form of the immunogen.

Essentially these results parallel our recent observations with a synthetic peptide derived from the HBsAg sequence 124 to 147 (peptide OS, Ref. 13). Because both the HBsAg and HIV-derived peptides are capable of spontaneous self-assembly on cleavage from the solid support, we were interested in determining whether co-oligomers of the HBsAg and HIV-1-derived sequences could be generated to give a single molecule that includes epitopes from both pathogens. This is schematically illustrated in Figure 5.

Equimolar amounts of resin containing either protected peptide HV or protected peptide OS were thoroughly mixed and subjected to the hydrogen fluoride cleavage procedure to achieve simultaneous side chain deprotection and cleavage from solid support. The resulting product was purified by cation exchange HPLC using a nonlinear NaCl gradient (Fig. 6). Two major peaks were determined which eluted at salt concentrations of 0.2 and 0.98 M, respectively (*peaks 1* and *2*, Fig. 6*A*). *Peak 2* was purified by preparative HPLC and its profile is shown in Figure 6*B*. An Ellman test (30) for purified *peak 2* (peptide OS-HV) was negative, indicative of quantitative disulfide bond formation.

To determine that peptide OS-HV contains both the OS and HV sequence covalently bound to each other, we employed a sandwich ELISA protocol. Wells of ELISA plates were first coated with mouse mAb against the "a" determinant of HBsAg (MAB A2-73, a kind gift from Dr. J. Pillot, Pasteur Institute, Paris). Antibody-coated wells were then incubated with peptide OS-HV to allow the

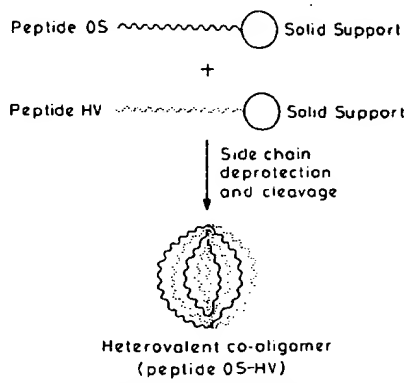

Figure 5. A schematic illustration that depicts the generation of the hybrid peptide OS-HV. For experimental details see *Materials and Methods*.

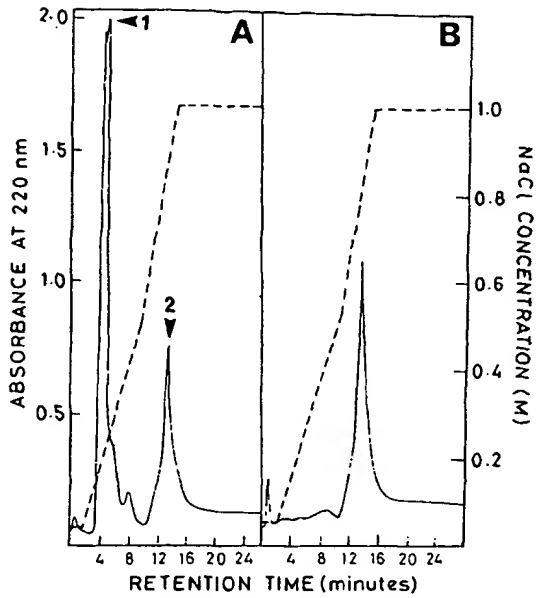

Figure 6. An HPLC profile of crude (*A*) and purified (*B*) preparations of peptide OS-HV. For experimental details, see *Materials and Methods*. The chart speed was 0.1 cm/min and the flow rate 1 ml/min. The wavelength of detection was 220 nm.

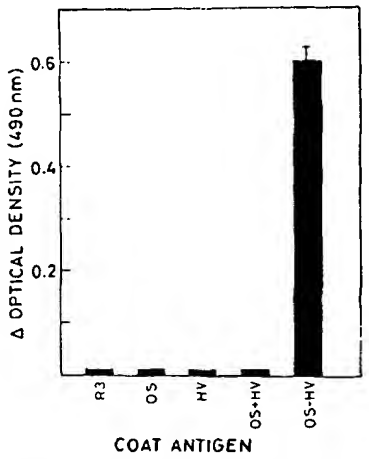

Figure 7. A sandwich ELISA assay to demonstrate that the OS and HV components are covalently linked in peptide OS-HV. For experimental details, refer to *Materials and Methods*. Values are the mean of triplicate determinations where the deviation from mean was less than 25%.

antibody to bind peptide through its OS sequence. Subsequently wells were washed and incubated with human anti-HIV-1-positive serum to allow for binding to the HV sequence in the peptide OS-HV. Bound human IgG was then detected by appropriately labeled second antibody (see *Materials and Methods*).

The results of the experiment are shown in Figure 7, where it can be seen that only peptide OS-HV gave a positive result. No signal was observed for either a physical mixture of peptides OS and HV or the individual components alone or for an irrelevant peptide included as an additional control. The results clearly demonstrate that purified *peak 2* in Figure 6 represents a co-oligomer of peptides OS and HV.

Conformational epitopes expressed by individual peptide OS and HV are retained in the hybrid construct OS-HV. In a previous communication, we have demonstrated that rabbit polyclonal sera obtained in response to peptide OS recognize only the disulfide bonded oligomeric form of peptide OS and not its reduced analog (13). Here we have shown that the anti-HV peptide response in rabbits is also conformation-specific (Fig. 4). Thus these different anti-peptide antisera should serve as useful probes to gauge the fidelity with which the epitope integrity of the individual component is retained in the hybrid OS-HV molecule. Figure 8*A* shows the results of an experiment when rabbit anti-HV antisera was titrated against a coat of either HV or OS-HV in an ELISA procedure. As is evident, both peptides showed comparable reactivity suggesting that the epitope expressed by peptide HV is not significantly altered when present as a component of the OS-HV molecule. When rabbit anti-OS antibodies were titrated against peptide OS or peptide OS-HV, near identical reactivities were again obtained (Fig. 8*B*).

Further, peptide OS-HV but not peptide HV was found to cross-react with polyclonal antibodies specific for HBsAg of either subtype ad or ay (Fig. 8*C*) indicating that the OS sequence in OS-HV also express the conformational and disulfide-dependent "a" epitope of HBsAg as has previously been shown for peptide OS (13).

These observations were further substantiated with human sera positive for either anti-HBsAg or anti-HIV antibodies. Table II shows the results obtained from a study of cross-reactivity of human anti-HBsAg sera with either peptide OS, peptide OS-HV or the linear, monomeric form of peptide OS (peptide MS). Comparable reactivity was obtained with peptide OS and OS-HV, whereas the peptide MS displayed poor antigenicity. Similar results were obtained with human anti-HIV antisera where peptide HV and OS-HV showed near identical antigenicity, whereas that of the linear, monomeric analog of peptide HV (peptide MHV) was low to negligible (Table III).

Cumulatively the results described in Figure 8 and Tables II and III clearly demonstrate that generation of the hybrid OS-HV molecule does not greatly perturb those conformational requirements necessary for optimal antigenic expression of the individual components.

Peptide OS-HV is immunogenic and elicits conformation specific antibodies against its individual components. Though peptide OS-HV contains both the OS- and HV-derived epitope, it was of interest to determine whether immunization with the hybrid would produce an antibody response against both components. To ascertain this, mice of four different H-2 haplotypes were immunized with either peptide OS, peptide HV, or peptide OS-HV and the resulting antisera were screened against either peptide OS or peptide HV. The results from this experiment are presented in Table IV.

Immunization with peptide OS-HV yielded an anti-OS and an anti-HV antibody response in all the four strains of mice studied (Table IV). Interestingly, although immunization with peptide HV alone gave no detectable anti-HV antibody response in B10.A mice, anti-HV antibodies were observed when OS-HV was used as immunogen. This suggests that the OS sequence in OS-HV provides help for production of an anti-HV response in B10.A mice.

The anti-OS titer obtained with peptide OS-HV was somewhat lower than that obtained on immunization with peptide OS alone, in all the four haplotypes, with B10.BR mice showing the most significant reduction (Table IV). On the other hand, anti-HV antibody titers ob-

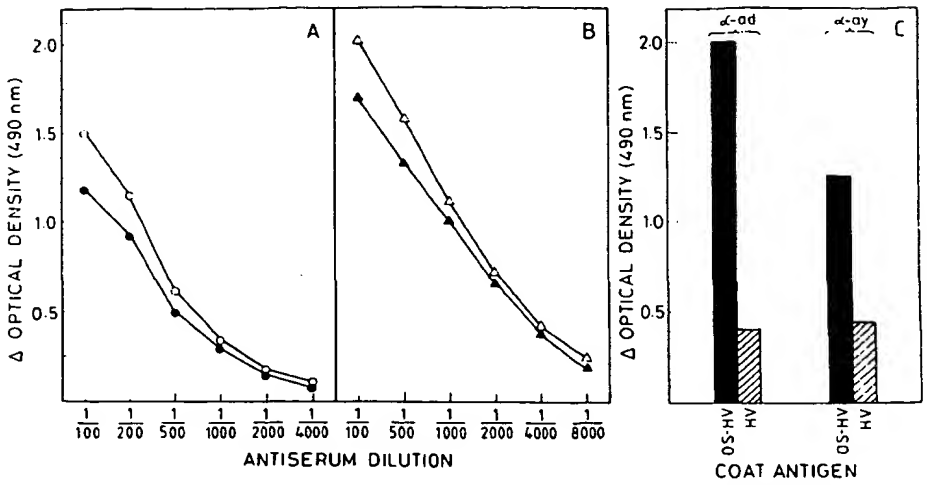

Figure 8. The conformational integrity of the OS and HV components is retained in the hybrid peptide OS-HV. (*A*) ELISA titration profile obtained with rabbit anti-OS antibodies (13) against a coat Ag of either peptide OS (○) or peptide OS-HV (●). (*B*) Titration profile obtained with rabbit anti-HV antibodies against either peptide HV (△) or peptide OS-HV (▲). (*C*) The reactivity obtained for polyclonal anti-HBsAg antisera specific for HBsAg of either subtype ad or ay against a coat of either peptide HV or OS-HV at an antiserum dilution of 1/100. Values are the absorbance obtained at 490 nm and for *panels A* and *B* backgorund absorbance obtained with preimmune serum at each dilution was subtracted. Each panel is a representative of three independent experiments.

TABLE II

Reactivity of human anti-HGsAg sera with peptides OS, OS-HV, and MS[a]

Sample	OD_{490} with Coat Peptides of		
	OS	OS-HV	MS
1	0.92	1.02	0.31
2	0.56	0.50	0.0
3	0.47	0.41	0.09
4	0.41	0.30	0.06
5	1.31	0.94	0.22
6	0.70	1.20	0.26
7	0.32	0.39	0.14
8	0.36	0.33	0.09
9	0.44	0.46	0.14
10	0.29	0.22	0.08

[a] Wells coated with either 1 μg of peptides OS or MS or with 2 μg of peptide OS-HV were incubated with 100 μ of a 1/10 dilution of sera from individuals who had received three doses of the Hepatitis B vaccine (see *Materials and Methods*). Values presented are the mean absorbance of duplicate samples obtained at 490 nm after subtracting the absorbance of a negative control serum sample also performed in duplicate.

TABLE III

Reactivity of human anti-HIV-1 antisera with peptides HV, OS-HV, and MHV[a]

Sample	OD_{490} with Coat Peptides of		
	HV	OS-HV	MHV
1	0.86	0.92	0.10
2	1.01	1.04	0.05
3	1.12	1.11	0.22
4	1.01	1.08	0.0
5	0.45	0.45	0.10
6	1.06	1.10	0.02
7	1.19	1.17	0.25
8	1.12	1.08	0.10
9	0.81	0.84	0.0
10	0.77	0.68	0.02
11	0.80	0.80	0.0
12	0.68	0.61	0.0
13	1.17	1.22	0.15
14	0.62	0.69	0.0
15	0.83	0.85	0.0
16	0.64	0.61	0.0

[a] The experimental protocol followed is identical to that described for Table II except that the coat antigens were either 1 μg/well of peptides HV or MHV and 2 μg/well of peptide OS-HV.

tained with either peptide HV or OS-HV were comparable in all the responder strains. Furthermore, the anti-OS and anti-HV antibodies, obtained in all the four mouse strains with peptide OS-HV were also conformation-specific as they were incapable of recognizing the linear, monomeric analogs of either peptide OS or peptide HV (Fig. 9).

The Th cell epitope resident in OS and HV sequences are codominant in the OS-HV hybrid peptide. The immunogenicity of peptides OS (13) and HV (Fig. 4) in the absence of a carrier protein indicates that both of these sequences include Th cell epitopes. We next investigated the status of these Th cell epitopes in the hybrid peptide OS-HV. Lymphocytes from the mouse groups of the four different strains that were primed with either peptide OS, peptide HV or peptide OS-HV (Table IV) were challenged in vitro with either peptide OS or HV and the stimulation index determined. These results are presented in Table V. A proliferative response was obtained to both peptide OS and HV in mice immunized with peptide OS-HV in all the four haplotypes demonstrating that the Th cell determinants in the component peptide retain their activity in the hybrid peptide OS-HV. Further the proliferative response against a given peptide was essentially the same in mice primed with either the individual peptide or the hybrid peptide OS-HV in most cases (Table IV). Thus the Th cell determinants resident in OS and HV sequences remain codominant in the hybrid peptide in three of the four mouse strains tested. The only exception to this was C57BL/6 mice where a high stimulation index was obtained with mice primed and challenged with peptide OS (Table IV). The unusually, high stimulation index observed for the group is somewhat surprising particularly because C57BL/6 mice do not represent the highest antibody responders to peptide OS (Table IV).

Another interesting finding in Table V is that lymphocytes from B10.A mice primed with peptide HV respond to an in vitro challenge with same peptide. This contrasts with the result in Table IV where the B10.A mice was found to be incapable of producing an anti-HV antibody response to peptide HV as immunogen. However, an anti-HV response was obtained when the immunogen was peptide OS-HV. These results indicate that the lack of an antibody response to peptide HV in B10.A mice is not due to "hole" in either the B or T cell repertoire. The reason for this observation is not clear at the present time and

TABLE IV

Murine humoral responses to individual components of peptide OS-HV on immunization with peptide OS-HV[a]

No.	Mouse Strain	H-2 Haplotype	Immunogen	1/Antibody Titer Against Peptide	
				OS	HV
1	B10.A	a	OS	8000	
			HV		ND[b]
			OS-HV	2000	2000
2	C57BL/6	b	OS	16000	
			HV		32000
			OS-HV	4000	64000
3	B10.M	f	OS	4000	
			HV		32000
			OS-HV	8000	64000
4	B10.BR	k	OS	32000	
			HV		2000
			OS-HV	1000	1000

[a] Groups of four mice from each strain were immunized with 20 μg of either peptide OS, peptide HV, or the hybrid peptide OS-HV (see *Materials and Methods*). Sera obtained from each group after a booster immunization were titrated against either peptide OS or peptide HV as indicated. Titers are expressed as that dilution of serum (1/1000 or greater) that gave an absorbance 2.1 times greater than that obtained for the preimmune serum at a constant dilution of 1/1000. Values are expressed as a reciprocal of antibody titers and are a representative of three independent experiments.
[b] ND, not detectable at a serum dilution of 1/100.

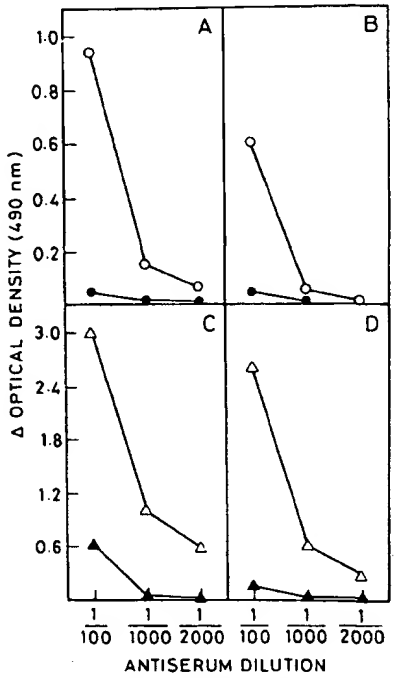

Figure 9. The hybrid peptide OS-HV elicits conformation-specific anti-OS and anti-HV antibody responses in a variety of mouse strains. Pooled sera (see *Materials and Methods*) obtained from B10.M (*A* and *C*) and C57BL/6 (*B* and *D*) were titrated against either peptide OS (○) or peptide MS (●) (*A* and *B*) at the indicated dilutions by ELISA. In a separate experiment, these sera were also titrated against either peptide HV (△) or peptide MHV (▲) (*C* and *D*). Values are the absorbance obtained at 490 nm after subtracting absorbance obtained for preimmune serum at each dilution. Though results from only two mouse strains are shown here, B10.A and B10.BR mice also gave essentially identical results.

TABLE V

Lymphocyte proliferation responses to component peptides in mice primed with either peptides OS, HV, or the hybrid peptide OS-HV[a]

Mouse strain	Immunogen	S.I. on Challenge with Peptide[b]	
		OS	HV
B10.A	OS-HV	55.2 ± 1.6	
	HV		11.8 ± 2.0
	OS-HV	5.2 ± 1.5	8.1 ± 1.1
C57BL/6	OS	5.3 ± 0.2	
	HV		4.2 ± 1.4
	OS-HV	5.7 ± 1.9	5.3 ± 1.2
B10.M	OS	5.9 ± 0.4	
	HV		5.2 ± 1.5
	OS-HV	4.6 ± 1.0	16.2 ± 3.7
B10.BR	OS	4.3 ± 0.6	
	HV		8.1 ± 2.3
	OS-HV	6.5 ± 2.2	8.3 ± 1.6

[a] For the experimental protocol (see *Materials and Methods*). Lymphocytes from mice primed with either peptide OS or HV were challenged in vivo with the homologous peptide whereas lymphocytes from mice primed with peptide OS-HV were challenged separately with peptides OS and HV.
[b] S.I., stimulation index. Values are mean (±SE) of S.I. obtained from three to five wells.

further experiments would be needed to unravel this. It is possible that the T cells activated by peptide HV are, for some reason, incapable of stimulating the anti-HV antibody producing B cells.

DISCUSSION

In a companion study (13) we showed that a synthetic peptide derived from the major protein of hepatitis B surface Ag could spontaneously self-assemble via disulfide-mediated oligomerization to reconstitute a conformational, group-specific determinant of HBsAg. The ability to derive such conformational epitope-mimic sequences from native Ag clearly opens up novel possibilities for developing peptide vaccines. The importance of these observations is further underscored by our more recent results that a myristylated form of this peptide is highly immunogenic in rhesus monkeys when alum is used as an adjuvant (manuscript submitted).

The present investigation was undertaken to determine whether these observations could be extended to other viral Ag. Given that a major drawback with synthetic peptide immunogens as candidate vaccines is their inability to accurately represent their conformationally restricted counterparts in the native protein (19, 31), this question assumes significance. A chimeric peptide that included sequences from both the envelope proteins of HIV-1 was shown to form disulfide-bonded multimers in a manner that was reminiscent of the HBsAg-derived peptide (13). Thus the oligomeric peptide, but not its linear monomeric analog, was recognized by human anti-HIV-1 antisera indicating that peptide HV represents a disulfide-induced conformational epitope on the native Ag. Further, peptide HV was found to be immunogenic in rabbits eliciting a predominant response against a conformational epitope that is either the same or overlaps with one that is commonly recognized on the native Ag when presented to the human immune system.

Taken together, the results obtained with peptide HV closely parallel those observed with the HBsAg derived peptide OS (124–147) (13) suggesting that it may have a more general applicability. Indeed as we have shown here, the ability to design a peptide with an inherent self-assembling capacity clearly enhances the scope of this approach and has important implications for the eventual development of successful peptide vaccines. It must be noted here that the design of peptide HV did not include the capacity to elicit a virus neutralizing antibody response. Rather this peptide was used as an exploratory model for future developmental work.

Because we had two separate peptides with intrinsic self-oligomerization properties, one derived from HBsAg and the other from HIV-1, the next question we asked was whether it would be possible to co-oligomerize these two sequences. In other words, could one generate a single molecule that encodes epitopes from both pathogens? Such a system might eventually prove useful in developing a polyvalent vaccine that can confer simultaneous protection against multiple pathogens which are transmitted by common routes and likely to lead to coinfection (e.g., HBV and HIV) (32).

By simply mixing equimolar amount of the independent peptide resins, we were able to obtain in situ, the hybrid peptide OS-HV upon cleavage. It was subsequently demonstrated that the hybrid peptide displayed the conformation dependent epitopes of both OS and HV. Indeed, at least at the gross level, the HBsAg and HIV related epitopes were expressed as efficiently in the hybrid peptide OS-HV as by the individual component peptides OS and HV, respectively. From this we infer that the parameters that influence the final conformation of OS and HV sequences in the oligomeric peptides OS and HV are not significantly affected when they are co-oligomerized to generate hybrid peptide OS-HV. Immunization with the peptide resulted in an antibody response against both components and these antibodies showed marked preference for the disulfide-bonded oligomeric forms of the OS and HV peptides. Thus it is possible to create polyvalent peptide molecules that include conformational epitopes from a variety of pathogens which in turn can elicit a conformation-specific humoral response against each of the individual component epitopes.

An analysis for the status of the OS and HV resident Th cell epitopes in the hybrid peptide OS-HV revealed that both the epitopes (or sets of epitopes) remain active in all the four strains of mice used for the study. Indeed in three of these strains the level of in vitro stimulation obtained on challenge was comparable regardless of whether the mice were primed with a component peptide alone or with the hybrid peptide OS-HV peptide, suggesting that the Th epitope contributed by the OS and HV sequences are codominant in the hybrid peptide. This is an important finding as it circumvents a primary obstacle in design and development of polyvalent candidate vaccines.

The traditional approach toward polyvalent immunogens has been to tandemly link characterized epitopes into an immunologically complex linear sequence by using either synthetic peptide (33, 34) or recombinant DNA (35) methodologies. However, this strategy suffers from the potential drawback of selective immunodominance which probably results from intramolecular competition between T cell epitopes (33, 36). Thus immunization with such linear constructs results in a hierarchy between the component epitopes so that the immune response is directed toward the more immunogenic epitopes, whereas the less immunogenic T cell epitopes remain cryptic (33, 36). It is not clear at present why peptide OS-HV does not show the phenomenon of intramolecular epitope competition and experiments are underway to address this. We believe that there are two factors which contribute toward the observed codominance of Th cell epitopes in this model. First the OS and HV sequences are linked to each other solely by disulfide bonds in peptide OS-HV. Second, this peptide being of an oligomeric nature is presumably large enough to enter the Ag processing pathway (37, 38). Thus on uptake by the APC a simple reduction of disulfide bonds in either the pre-lysosomal or lysosomal compartments (39) is sufficient to release the individual OS and HV fragments which can display their full immunogenicity. However, this is only a hypothesis and needs to be verified.

Although the term polyvalent has often been used to denote a designed immunogen incorporating a multiplicity of epitopes (40), we propose that classification of these as either multi- or heterovalent would be more descriptive. Thus a multivalent immunogen can be one where its component epitopes were derived from one or more proteins of a single pathogen (e.g., Ref. 41). On the other hand the term heterovalent can be applied to such immunogens where their component epitopes were derived from more than one pathogen (e.g., peptide OS-HV). The disulfide-mediated approach has been used in the past to generate a synthetic hybrid peptide containing selected epitopes from the merozoite specific protein of *Plasmodium falciparum* (41). However, in this instance, the question of conformational B cell determinants was not addressed and neither was the relative immunodominance of the component epitopes. Clearly then, the methodology involving generation of self-assembling conformational epitope-mimic sequences as discussed in this and the previous report (13) provides a powerful and versatile new strategy for developing both multi- and heterovalent immunogens that may eventually find success as vaccines.

Perhaps the most important and immediate application of our system is that it permits one to address the problem of antigenic variation. This would be particularly relevant to the problems associated with developing a vaccine against HIV-1. The principal neutralizing determinant on gp120 is localized to a region between a disulfide loop spanning from residues 303 and 338 (42–44). However, this sequence represents a hypervariable region of the protein and consequently antibodies raised against it have only type specific virus neutralizing ability (44, 45). We have recently constructed a multivalent co-oligomer molecule that includes conformationally restricted principal neutralizing determinant sequences from five distinct isolates of HIV-1. The immunologic properties of this immunogen and the ability of antibodies raised against it to neutralize a broader cross-section of HIV-1 isolates is currently under investigation.

Acknowledgment. The authors wish to thank Ms. Radha for her excellent secretarial assistance.

REFERENCES

1. Arnon, R. 1980. Chemically defined antiviral vaccines. *Annu. Rev. Microbiol. 34:593.*
2. Braciale, T. J., L. R. Morrison, M. T. Sweetser, J. Sambrook, M-J. Getlung, and V. L. Braciale. 1987. Antigen presenting pathways to class I and class II MHC-restricted T lymphocytes. *Immunol. Rev. 98:95.*
3. Buus, S., A. Sette, and H. M. Grey. 1987. The interaction between protein-derived immunogenic peptides and Ia. *Immunol. Rev. 98:115.*
4. Kourilsky, P., and J-M. Claverie. 1989. MHC-antigen interaction: what does the T cell receptor see? *Adv. Immunol. 45:107.*
5. Panina-Bordignon, A. Tan, A. Ternijtelen, S. Demotx, G. Corradin, and A. Lanzavecchia. 1989. Universally immunogenic T cell epitopes: Promiscuous binding to the MHC class II and promiscuous recognition by T cells. *Eur. J. Immunol. 19:2237.*
6. Etlinger, H. M., D. Gillessen, H. W. Lahm, H. Matile, H. J. Schonfeld, and A. Trzeciak. 1990. Use of prior vaccination for the development of new vaccines. *Science 249:423.*
7. Sinigaglia, F., M. Guttinger, J. Kligus, D. M. Doran, H. Matile, H. Etlinger, A. Trzeciak, D. Gillesen, and J. R. L. Pink. 1988. A malaria T cell epitope recognized in association with most mouse and human MHC class II molecules. *Nature 336:778.*
8. Morein, B., and K. Simons. 1985. Subunit vaccines against enveloped viruses: virosomes, micelles and other protein complexes. *Vaccine 3:83.*
9. Boudrealt, A., and L. Thibodeau. 1985. Mouse response to influenza immunosomes. *Vaccine 3:231.*
10. Goodman-Snitkoff, G., L. E. Eisele, E. P. Heimer, A. M. Felix, T. T. Andersen, T. R. Fuerst, and R. J. Mannino. 1990. Defining minimal requirements for antibody production to peptide antigens. *Vaccine 8:257.*
11. Van Regenmortel, M. H. V. 1990. The structure of viral epitopes. In *Immunodiagnosis of Virus. II. The Basis for Serodiagnosis and Vaccines.* M. H. V. Van Regonmortel and A. R. Neurath, eds. Elsevier Science Publishers, The Netherlands, p. 1.
12. Berman, P. W., T. J. Gregory, L. Riddle, G. R. Nakamura, M. A. Champe, J. P. Porter, F. M. Wurn, R. D. Hershberg, E. K. Cobb, and J. W. Eichberg. 1990. Protection of chimpanzees from infection by HIV-1 after vaccination with recombinant glycoprotein gp120 but not gp160. *Nature 345:622.*
13. Manivel, V., R. Ramesh, S. K. Panda, and K. V. S. Rao. A synthetic peptide spontaneously self-assembles to reconstruct a group specific, conformational determinant of hepatitis B surface antigen. *J. Immunol. 148:4006.*
14. Rossman, M. G. 1989. The canyon hypothesis. Hiding the host cell receptor on a viral surface from immune surveillance. *J. Biol. Chem. 264:14587.*
15. Klenk, H. D. 1990. Influence of glycosylation on antigenicity of viral proteins. In *Immunochemistry of Viruses. II. The Basis for Serodiagnosis and Vaccines.* M. H. V. Van Regenmortel and A. R. Neurath. eds. Elsevier Science Publishers, The Netherlands, p. 25.
16. Laver, W. G., G. M. Air, R. G. Webster, and S. S. Smith-Gill. 1990. Epitopes on protein antigens: misconception and realities. *Cell 61:553.*
17. Geysen, H. M., T. J. Mason, and S. J. Rodda. 1989. Cognitive features of continuous antigenic determinants. In *Synthetic Peptides: Approaches to Biological Problems.* J. P. Tam and E. T. Kaiser, eds. Alan R. Liss, New York, p. 19.
18. Colman, P. M. 1988. Structure of antibody-antigen complexes: implications for immune recognition. *Adv. Immunol. 43:99.*
19. Bidart, J. M., F. Troalen, P. Ghillani, N. Rouas, A. Razafindratsita, C. Bohoun, and D. Bellet. 1990. Peptide immunogen mimicry of a protein-specific stuctural epitope on human chorionogonadotropin. *Science 248:736.*
20. Neurath, A. R., S. B. H. Kent, and N. Strick. 1982. Specificities of antibodies elicited by a synthetic peptide having a sequence in common with a fragment of a virus protein, the hepatitis B surface antigen. *Proc. Natl. Acad. Sci. USA 79:7871.*
21. Howard, C., H. Strick, A. Buckley, S. Brown, and M. Steward. 1989. The use of synthetic peptides for the analysis of the envelope of the hepatitis B virus (HBV). In *Synthetic Peptides: Approaches to Biological Problems.* J. P. Tam and E. T. Kaiser, eds. Alan R. Liss, New York, p. 211.
22. Dreesman, G. R., Y. Sanchez, I. Ionescu-matiu, J. P. Sparrow, H. R. Six, D. L. Peterson, F. B. Hollinger, and J. L. Melnick. 1982. Antibody to hepatitis B surface antigen after a single inoculation of uncoupled synthetic hepatitis B surface antigen peptides. *Nature 295:158.*
23. Ionescu-Matiu, I., R. C. Kennedy, J. P. Sparrow, A. R. Culwell, Y. Sanchez, J. L. Melnick, and G. R. Breesman. 1983. Epitopes associated with a synthetic hepatitis B surface antigen peptide. *J. Immunol. 130:1947.*
24. Bhatnagar, P. K., E. Papas, H. E. Blum, D. R. Milich, D. Nitecki, M. J. Karels, and G. N. Vyas. 1982. Immune response to synthetic peptides analogues of hepatitis B surface antigen specific for the "a" dterminant. *Proc. Natl. Acad. Sci. USA 79:4400.*
25. Brown, S. E., C. R. Howard, A. J. Zuckermann, and M. W. Steward. 1984. Determiantion of affinity of antibodies to hepatitis B surface antigen in human sera. *J. Immunol. Methods 72:41.*
26. Wampler, D. E., E. D. Lehman, J. Boger, W. J. McAleer, and E. J. Scolnick. 1985. Multiple chemical forms of hepatitis B surface antigen produced in yeast. *Proc. Natl. Acad. Sci. USA 82:6830.*
27. Gnann, J. W., J. A. Nelson, and M. B. A. Oldstone. 1987. Fine mapping of an immunodominant domain in the transmembrane glycoprotein of human immunodeficiency virus. *J. Virol. 61:2639.*
28. Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staahl, and H. Wolf. 1987. Computer assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *J. Virol. 61:570.*
29. Chou, P. Y., and G. D. Fasman. 1979. Prediction of beta turns. *Biophys. J. 26:367.*
30. Ellman, G. L. 1959. Tissue sulfhydryl groups. *Arch. Biochem. Biophys. 82:70.*
31. Brigido, M. M., J. Sabbaga, and R. R. Brentani. 1990. Are synthetic peptides suitable for the detection of continuous epitopes only? *Immunol. Lett. 24:191.*
32. Rosenberg, Z. F., and A. S. Fauci. 1990. The immunopathogenesis of HIV infection. *Adv. Immunol. 47:377.*
33. Ria, F., B. M. C. Chan, M. T. Scherer, J. A. Smith, and M. L. Gefter. 1990. Immunological activity of covalently linked T-cell epitopes. *Nature 343:381.*
34. Golvano, J., J. J. Lasarte, P. Sarobe, A. Gullon, J. Prieto, and F. Borras-Cuesta. 1990. Polarity of immunogens: Implications for vaccine design. *Eur. J. Immunol. 20:2363.*
35. Kumar, V., V. J. Bansal, K. V. S. Rao, and S. Jameel. 1992. Gene assembly and *E. coli* expression of an immunologically reactive novel multiple epitope polypeptide (MEP-1): hepatitis B virus envelope protein epitopes. *Gene 110:137.*
36. Perkins, D. L., G. Berriz, T. Kamradt, J. A. Smith, and M. L. Gefter. 1991. Immunodominance: intramolecular competition between T cell epitopes. *J. Immunol. 146:2137.*
37. Long, E. O., and S. Jacobson. 1989. Pathways of antigen processing and presentation in CTL: defined by mode of virus entry? *Immunol. Today 10:45.*
38. Smith, J. A. 1989. Synthetic peptides: tools for elucidating mechanisms of protein antigen processing and presentation. In *Synthetic Peptides: Approaches to Biological Problems.* J. P. Tam and E. T. Kaiser, eds. Alan R. Liss, New York, p. 31.
39. Harding, C. V., D. S. Collins, J. W. Slot, H. J. Geuze, and E. R. Unanue. 1991. Liposomes-encapsulated antigens are processed in lysozomes, recycled and presented to T cells. *Cell 64:393.*
40. Jolivet, M., L. Lise, H. Grasse-Masse, A. Tartar, F. Audibert, and L. Chedid. 1990. Polyvalent synthetic vaccines: relationship between T epitopes and immunogenicity. *Vaccine 8:35.*
41. Patarayyo, M. E., R. Amador, P. Clavigo, A. Moreno, F. Guzman, P. Romero, R. Tascon, A. Franco, L. A. Murillio, G. Ponton, and G. Trujillo. 1988. A synthetic vaccine protects humans against challenge with asexual stages of *Plasmodium falciparum* malaria. *Nature 332:158.*
42. Rusche, J. R., K. Jahaverian, C. McDanal, J. Petro, D. L. Lynn, R. Grimaila, A. Langlois, R. C. Gallo, L. O. Arthur, P. J. Fischinger, D. P. bolognesi, S. D. Putney, and T. J. Matthews. 1988. Antibodies that inhibit fusion of human immunodeficiency virus-infected cells bind to a 24 amino acid sequence of the viral envelope gp120. *Proc. Natl. Acad. Sci. USA 85:3198.*
43. Matsushita, S., M. Robert-Guroff, J. Rusche, A. Koito, T. Hattori, H. Hoshino, K. Jahaverian, K. Takatsui, and S. D. Putney. 1988. Characterization of a human immunodeficiency virus neutralizing monoclonal antibody and mapping of the neutralizing epitope. *J. Virol. 62:2107.*
44. Jahaverian, K., A. J. L. Anglois, C. McDanal, K. L. Ross, L. I. Eckler, C. L. Jellis, A. T. Profy, J. R. Rusche, D. P. Bolognesi, S. D. Putney, and T. J. Matthews. 1989. Principal neutralizing domain of the human immunodeficienty virus type I envelope protein. *Proc. Natl. Acad. Sci. USA 86:6768.*
45. Palker, T. J., M. E. Clark, A. G. Langlois, T. J. Matthews, K. J. Weinhold, R. R. Randall, D. P. Bolognesi, and B. F. Haynes. 1988. Type specific neutralization of human immunodeficiency virus with antibodies to env-coded peptides. *Proc. Natl. Acad. Sci. USA 85:1932.*

0022-1767/92/14812-4006$02.00/0
THE JOURNAL OF IMMUNOLOGY

Vol. 148, 4006–4011, No. 12, June 15, 1992
Printed in U.S.A.

A SYNTHETIC PEPTIDE SPONTANEOUSLY SELF-ASSEMBLES TO RECONSTRUCT A GROUP-SPECIFIC, CONFORMATIONAL DETERMINANT OF HEPATITIS B SURFACE ANTIGEN

VENKATASAMY MANIVEL,* RAJAGOPAL RAMESH,† SUBRAT K. PANDA,† AND KANURY V. S. RAO[1]*

*From the *Virology Group, International Centre for Genetic Engineering and Biotechnology, NII Campus, Shaheed Jeet Singh Marg, New Delhi-110 067, India, and †Department of Pathology, All India Institute of Medical Sciences, New Delhi-110 029, India*

A cysteine-rich peptide of sequence 124 to 147 of the major protein of hepatitis B surface Ag (HBsAg) was synthesized. On cleavage and subsequent work-up it was found that all of the cysteine sulfhydryl groups had spontaneously formed disulfide bonds to yield a heterogenous mixture of multiple forms with molecular masses ranging from 8 to 35 kDa (peptide OS[124–147]). In a direct ELISA peptide OS[124–147] showed a high degree of cross-reactivity with polyclonal anti-HBsAg antiserum whereas the HBsAg-related antigenicity of its disulfide-reduced analogs was insignificant. Peptide OS[124–147] was also recognized by all 15 of the anti-HBsAg-positive human sera tested. Further studies revealed that peptide OS[124–147] represents the conformational, disulfide-dependent "a" determinant of HBsAg and elicits antibodies that cross-react with a variety of HBsAg subtypes. Anti-peptide antibodies bound to the corresponding native epitope with an apparent affinity higher than that of homologous antisera. Finally, polyclonal anti-OS[124–147] antibodies could also immunoprecipitate purified Dane particles in solution. Together these studies indicate that peptide OS[124–147] represents an excellent candidate component of a peptide-based vaccine for hepatitis B.

The envelope of HBV2 consists of several protein species (see Refs. 1–3 for a review). The major component of the HBsAg is the S protein which is composed of 226 amino acids and is either glycosylated (gp 27) or nonglycosylated (p 24). In addition to the S protein lesser amounts of higher molecular weight proteins have also been detected (3, 4). Of these the middle (M) protein corresponds to the S protein with an additional 55 amino acids derived from the pre-S[2] component of the envelope (env) gene and is also present in two forms (gp33 and gp36) according to the extent of glycosylation. The large (L) protein is encoded by the pre-S1 region, the pre-S2 region and S gene and is present in a glycosylated (gp42) and a nonglycosylated (p39) form. This protein is of variable length according to the subtype (3).The possibility of a synthetic peptide vaccine for HBV had been suggested earlier (5, 6) and subsequent cloning of the HBV genome along with the deduced amino acid sequence (e.g., see Ref. 7) spurred a lot of interest in this area. Synthetic peptides have been successfully used to localize immunologically and functionally important determinants on the pre-S region of HBsAg (3, 8, 9). Further pre-S1 and pre-S2 derived peptides capable of inducing virus neutralizing antibodies have also been described (10–12). However, the S protein has proved more intractable to such analysis primarily because a majority of its antigenic determinants are dependent on conformation which in turn is induced by disulfide bond formation (13–15). Disulfide bonds also appear to be involved in intermolecular cross-linking of monomer S protein molecules (see Ref. 16), raising the possibility that the S protein may also express antigenic determinants dependent upon its quarternary structure.

Serologically HBsAg has one group-specific antigenic determinant "a" and two sets of mutually exclusive determinants "d" or "y" and "w" or "r" resulting in four major serotypes: adw, adr, ayw, and ayr (17). Using a panel of S protein-derived synthetic peptides Gerin et al. (18) observed that the group-specific "a" Ag was composed of at least three nonoverlapping sequences, whereas the sequence between residues 110 and 137 specified the major d/y subtype system. Bhatnagar et al. (19) have shown that the nonapeptide sequence 139 to 147 represents an essential part of the "a" determinant, though the presence of an additional "d" epitope within this sequence has also been suggested (20). The "a" epitope contained between residues 139 and 147 appears to be conformational because a cyclic form of the corresponding synthetic peptide was recognized by human anti-HBsAg antisera with higher affinity than the linear analog (21, 22).

An additional conformation-dependent "a" epitope has also been identified with a disulfide induced cyclic peptide corresponding to residues 122 to 137 of HBsAg (23, 24). In addition to the cyclic "a" epitope a sequential "y"

Received for publication December 9, 1991.
Accepted for publication March 20, 1992.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Address correspondence and reprint requests to Dr. Kanury V. S. Rao, Virology Group, International Centre for Genetic Engineering & Biotechnology, NII Campus, Shaheed Jeet Singh Marg, New Delhi-110 067, India.

[2] Abbreviations used in this paper: HBsAg, hepatitis B surface Ag; HBV, hepatitis B virus; RT, room temperature.

epitope has also been identified on this sequence (24). More recent studies indicate that the cyclic 139–147 peptide represents the more immunodominant epitope than the cyclic 122–137 peptide in humans vaccinated with either the plasma or recombinant-derived hepatitis B vaccine (25).

In this communication we report that a synthetic peptide corresponding to residues 124 to 147 of the S protein spontaneously oligomerizes to reconstitute a conformational group-specific antigenic determinant of HBsAg that is stringently dependent upon the integrity of the oligomer. Furthermore the oligomeric peptide is immunogenic and elicits high affinity antibodies capable of recognizing a variety of HBsAg subtypes and immunoprecipitating Dane particles.

MATERIALS AND METHODS

Materials. Plasma-derived HBsAg used in these experiments was the vaccine preparation (H-B-VAX) marketed by Merck Sharp and Dohme (Westpoint, PA). This preparation contains purified HBsAg at a protein concentration of 20 μg/ml. Polyclonal horse anti-HBsAg antiserum was purchased from Wellcome Laboratories (Beckenham, UK). HBsAg subtype-specific antisera used were the WHO International Reference reagents obtained from the Central Laboratory of the Netherlands Red Cross Blood Transfusion Service (Amsterdam, The Netherlands). HBsAg subtype standards used were either the WHO (subtypes adw, ayw) or NIH (subtypes ayr, ayw2, a2dw, and adw2) reference reagents and were a kind gift from Dr. M. S. Rajagopalan (National Institute of Immunology, New Delhi, India). All secondary antibodies conjugated to horseradish peroxidase were purchased from Sigma Chemical Co. (St. Louis, MO).

Peptide synthesis. Peptide S[124–147] (sequence, CTTPAQGNSMFPSCCCTKPTDGNC) was synthesized by the solid phase method (26) on 4-methylbenzhydrylamine resin using an Applied Biosystems model 430A automated peptide synthesizer. A 500-mg sample of the peptide-resin was cleaved and deprotected in the presence of 0.5 ml thioanisole with 8 ml of a 10% (v/v) solution of trifluoromethanesulfonic acid in trifluoroacetic acid at RT for 90 min. The crude product was precipitated in ether and filtered. The residue was dissolved in 10% acetic acid and lyophilized. Product obtained was dissolved in 15 ml of 5% acetic acid and dialyzed in a tubing with a molecular mass cut off of 1200 Da. (Sigma Chemical Co., St. Louis, MO) over 20 h against 5% acetic acid with four changes of 2 liters each. The dialyzed solution was lyophilized again to yield 70 mg of peptide OS[124–147] which was characterized by amino acid analysis. Both partially reduced (peptide TS[124–147]) and monomer (peptide MS[124–14]) derivatives were obtained by reduction with a 10-fold molar excess of dithiothreitol (RT, 3 h) followed by treatment with a 24-fold excess of iodoacetamide (30 min, RT). Peptides TS[124–147] and MS[124–147] could be purified by reverse phase HPLC on a C-18 column (μ Bondapak, 7.8 mm × 30 cm) using an aqueous gradient of 0 to 60% acetonitrile in 0.1% TFA over 40 min. Though peptides OS[124–147] and TS[124–147] could be visualized by silver staining after gel electrophoresis under nonreducing conditions, peptide MS[124–147] was not detectable. However, the latter peptide could be observed at the expected molecular mass (~3 kDa) by prior dansylation and subsequent exposure of the gel to UV light.

ELISA. Wells were coated with 1 μg of Ag in 100 μl of PBS (pH 7.4) at 37°C for 3 h. Subsequently they were blocked with 100 μl of 5% (w/v) solution of nonfat dry milk powder at 37°C for 2 h and washed thoroughly. Aliquots of 100 μl of indicated dilutions of antibody were added and incubated overnight at 4°C. After washing, wells were incubated with 100 μl of horseradish peroxidase-labeled secondary antibody at 37°C for 1.5 h. Chromogen used was *o*-phenylenediamine and the absorbance was measured at 490 nm.

Immunization with peptide OS[124–147]. A group of four BALB/c mice were immunized i.p. with 20 μg each of peptide OS[124–147] in CFA. Three weeks later (day 21) these animals were boosted with an identical dose of peptide in IFA. Mice were bled from the retroorbital plexus 7 days after the boost (day 28) and the sera pooled.

Two NZW rabbits were immunized i.m. with 200 μg each of peptide in CFA. Eight weeks later (day 56) the rabbits were boosted with an identical dose of peptide in IFA. Blood was collected from the retroorbital plexus on day 63. Assays of individual sera indicated that, although one of the rabbits responded fairly well to peptide, the other gave a relatively weaker response (~4-fold lower titer against peptide). Only the serum obtained from the higher responder rabbit was used for experiments described in this report.

Purification of Dane particles. Dane particles were purified from human plasma positive for HBV DNA by a previously described method (27). Presence of virions in the purified preparation was confirmed by transmission electron microscopy after negative staining with 2% phosphotungstic acid (pH 6.8) on a CM 10 Phillips microscope. Presence of HBV DNA in the preparation was also established by spot hybridization (28). The number of Dane particles was quantitated by calibration against a known standard quantity of cloned HBV DNA assuming one genome (3.2 kb) per Dane particle.

Immunoprecipitation of Dane particles by anti-peptide antiserum. Dane particles corresponding to about 40 pg DNA were diluted to 100 μl in PBS to which was added 10 μl of either rabbit anti-OS[124–147] serum, rabbit preimmune serum, or rabbit anti-HBsAg serum as positive control. The total volume was made up to 200 μl in PBS. This was incubated first at 37°C for 1 h and then at 4°C overnight. An aliquot of 100 μl of a protein A-Sepharose 4B suspension in PBS (Pharmacia, Sweden) was added and incubation continued with gently shaking at 37°C for 1 h followed by a 3-h incubation at 4°C. The suspension was then centrifuged at 2000 rpm for 15 min and supernatant collected. The protein A-Sepharose pellet was washed thrice with 3 ml each of PBS and bound material eluted with 100 μl of 0.1 M glycine-HCl buffer (pH 2.8). Both supernatant and eluate were processed for spot hybridization using previously described procedure (28).

RESULTS

Choice and synthesis of peptide S[124–147]. We chose to synthesize a peptide corresponding to residues 124 to 147 of the S protein because it was expected to contain the group-specific determinants localized to both 122–137 (23, 24) and 139–147 (19) segments. The sequence selected was that present in HBsAg of subtype adw (19) and contains five cysteine residues, one each at the amino and carboxyl termini with the remainder occurring at positions 137, 138, and 139.

Following synthesis and cleavage from resin an Ellman test (29) revealed that less than 2% of the cysteines bore free sulfhydryl groups; indicative of spontaneous and quantitative disulfide bond formation. Given the odd number of cysteine residues it was expected that this product (peptide OS[124–147]) would represent an oligomeric form. This was indeed found to be the case by gel electrophoresis which showed that peptide OS[124–147] was a mixture of multiple forms with apparent molecular masses ranging from about 8 to 35 kDa though a prominent band at a molecular mass of about 12 kDa was also present (Fig. 1). From the theoretical molecular mass of the monomer (2.5 kDa) we expect the latter band to correspond to a tetrameric or pentameric species. On reduction and subsequent carboxymethylation (*Materials and Methods*) two peaks could be purified by HPLC. The major peak corresponded to the monomer (peptide MS[124–147]) and a minor peak which by gel electrophoresis appears to represent a partially reduced trimeric species (peptide TS[124–147]) (Fig. 1).

Peptide OS[124–147] contains HBsAg-related antigenic determinants. All of the three peptides described above were tested for cross-reactivity with polyclonal anti-HBsAg antiserum by direct ELISA (*Materials and Methods*). Although peptide OS[124–147] displayed significant cross-reactivity, the HBsAg-related antigenicity of both peptides TS[124–147] and MS[124–147] was drastically reduced (Fig. 2). Specificity of antibody and OS[124–147] interaction could be demonstrated by dose-dependent competitive inhibition in the presence of increasing amounts of HBsAg (results not shown). Thus the peptide S[124–147] appears to spontaneously reconstitute one or more antigenic determinants related to

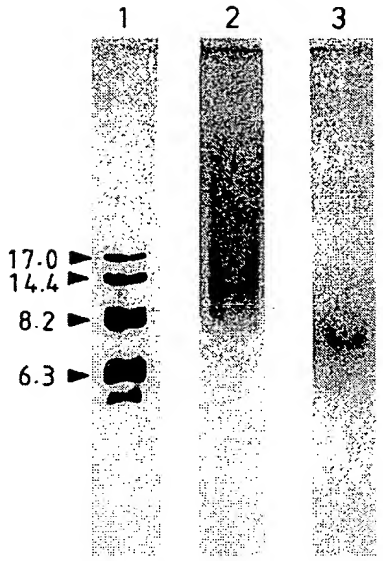

Figure 1. PAGE of peptides OS[124–147] and TS[124–147] under non-reducing conditions. Two micrograms of each peptide were dissolved in an equal volume of 2-ME-deficient sample buffer (8% (w/v) SDS, 0.25 M Tris, 40% (v/v) glycerol, and 0.005% bromphenol blue) and resolved on an 18% polyacrylamide gel. Visualization was by silver staining. Lanes: *1*, MW markers (in kDa); 2, peptide OS[124–147]; and *3*, peptide TS[124–147].

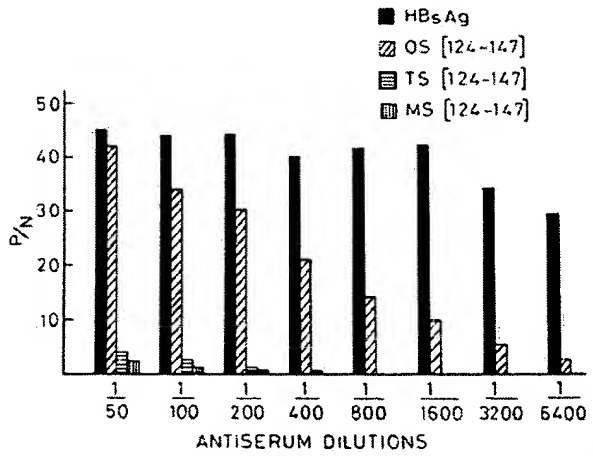

Figure 2. Polyclonal anti-HBsAg antiserum recognizes peptide OS[124–147] in an ELISA assay. For the experimental protocol see *Materials and Methods*. Figure shows a representative of three separate experiments where each determination was done in duplicate. Data are expressed as P/N ratios where N represents the absorbance obtained with an irrelevant peptide corresponding to the human IL-1β sequence 161 to 171 at a dilution of 1/100. Though the crude peptide OS[124–147] obtained as described in *Materials and Methods* was used for this experiment, similar results were obtained with peptide that was passed either through a gel-permeation or reverse phase HPLC column. Further a second batch of peptide OS[124–147] that was obtained by the hydrogen fluoride procedure of cleavage from the solid support (done by Multiple Peptide Systems, San Diego, CA) also gave identical results.

HBsAg which is dependent upon integrity of the disulfide bonds in the oligomer. The higher titer obtained with HBsAg (Fig. 2) indicates that peptide OS[124–147] does not represent all of the epitopes expressed by the S protein.

Peptide OS[124–147] contains the "a" determinant of HBsAg. Standard HBsAg subtype-specific antisera were screened for cross-reactivity with peptides OS[124–147], TS[124–147], and MS[124–147]. Significant reactivity was observed to peptide OS[124–147] with all three subtype-specific antisera though the corresponding antigenicity of peptides TS[124–147] and MS[124–147] was drastically reduced (Table I). This demonstrates that the epitope(s) expressed by peptide OS[124–147] represents the group specific "a" determinant of HBsAg (17) and that this "a" antigenicity is disulfide bond dependent.

Peptide OS[124–147] represents a dominant constituent of the HBsAg epitope repertoire in humans. To ascertain whether the epitope contained within the 124 to 147 sequence of HBsAg is also expressed in humans, anti-HBsAg antisera from 15 donors were titrated against either HBsAg or peptide OS[124–147]. Of these, 13 were individuals who had received either one, two, or three doses of a plasma-derived hepatitis B vaccine whereas the remaining two were from individuals who had previously recovered from a hepatitis B viral infection.

Table II gives the results of this experiment. All samples were positive for both anti-HBsAg and anti-OS[124–147] with titers being comparable in many cases. Sera negative for anti-HBsAg by conventional assays from four donors were also titrated in parallel. All of these were also negative in our experiment (results not shown), indicative of the specificity of the assay system. These results suggest that the HBsAg epitope(s) related to that of peptide OS[124–147] is dominant when presented to the human immune system.

It is not clear at present why the anti-OS[124–147] titer in some cases is greater than the anti-HBsAg titer (Table 2), though there is precedence for such observation (25).

TABLE I

Peptide OS[124–147] contains one or more conformation-dependent "a" determinants[a]

Subtype Specificity of Antiserum	P/N		
	OS[124–147]	TS[124–147]	MS[124–147]
ad	64.5	3.2	0.3
ay	54.0	5.4	0.5
ar	47.8	4.3	0.6

[a] Subtype-specific antisera (dilution, 1/100) used were the WHO International Reference reagents obtained from the Central Laboratory of the Netherlands Red Cross Blood Transfusion Services. Values are the mean of triplicate determinations.

TABLE II

Human anti-HBsAg antisera cross-react with peptide OS[124–147][a]

Sample No.	1/Antibody Titer	
	Anti-HBsAg	Anti-OS[124–147]
1	660	440
2	210	130
3	490	240
4	900	780
5	670	880
6	3620	1130
7	3050	2190
8	150	140
9	410	240
10	710	520
11	1160	440
12	1710	1360
13	4570	3430
14	150	530
15	2660	2040

[a] Wells coated with either peptide OS[124–147] or HBsAg were incubated with 100 μl of varying dilutions of sera from individuals either vaccinated with a plasma-derived hepatitis B vaccine (Nos. 1–13) or individuals with a prior clinical history of hepatitis B (Nos. 14 and 15). For normal serum the negative control serum provided with the AUSAB kit (Abbot Laboratories, Chicago, IL) was used. Total human Ig bound was measured and titers are expressed as that dilution of serum that gave an absorbance 2.1 times greater than that obtained for the normal serum control at the same dilution. Values are the mean of three independent determinations where each determination was done in duplicate.

Peptide OS[124–147] elicits antibodies cross-reactive with native HBsAg. To further examine the HBsAg-related epitope encoded within peptide OS[124–147] polyclonal anti-peptide antisera were raised in both mice and rabbits. Peptide OS[124–147] was found to be immunogenic in the absence of a carrier protein in both hosts and a primary IgG response could be boosted upon subsequent challenge with peptide. These antisera were titrated against either peptide or HBsAg and the results are shown in Figure 3. In both cases the titration profiles obtained against peptide and HBsAg were essentially similar indicating that at least a major proportion of antibodies elicited against peptide also cross-react with native HBsAg. This implies that the determinant(s) expressed by peptide OS[124–147] mimic the corresponding epitopes of HBsAg with high fidelity.

The majority of antibodies obtained in response to peptide OS[124–147] appear to be directed against one or more conformational determinants although very weak (<10%) to no cross-reactivity was observed with either peptide MS[124–147], peptide TS[124–147] or HBsAg reduced with 2-ME (results not shown).

In double antibody competition experiments it was found that binding of polyclonal anti-HBsAg antiserum to peptide OS[124–147] was inhibited in a dose-dependent manner by increasing concentrations of rabbit anti-peptide antiserum (Fig. 4, O). On the other hand, binding of rabbit anti-peptide antisera to native HBsAg was unaffected by increasing concentrations of polyclonal anti-HBsAg antisera (Fig. 4, ●) with only about 20% inhibition at a competitor dilution of 1:2 (not shown). Similarly anti-HBsAg antibodies were also unable to compete with anti-peptide antisera for peptide OS[124–147] (results not shown). In all cases tracer antibody was used at a constant dilution that was fivefold less than its titer for peptide OS[124–147] so as to maintain comparable levels of anti-peptide antibodies in both sets. Essentially identical results were obtained when a preparation of mouse polyclonal anti-HBsAg antiserum was used instead of the horse polyclonal anti-HBsAg preparation as described in Figure 4 (not shown).

Collectively these data indicate that, although polyclonal anti-HBsAg antibodies and anti-peptide antibodies recognize either a common or overlapping determinant on peptide OS[124–147], anti-peptide antibodies bind the corresponding native epitope on HBsAg with a much higher apparent affinity than do antibodies from the polyclonal anti-HBsAg antiserum preparation. It must be noted here that the antibody titer against homologous Ag was at least 10-fold higher in the case of the anti-HBsAg preparations when compared with the rabbit anti-peptide antiserum preparation.

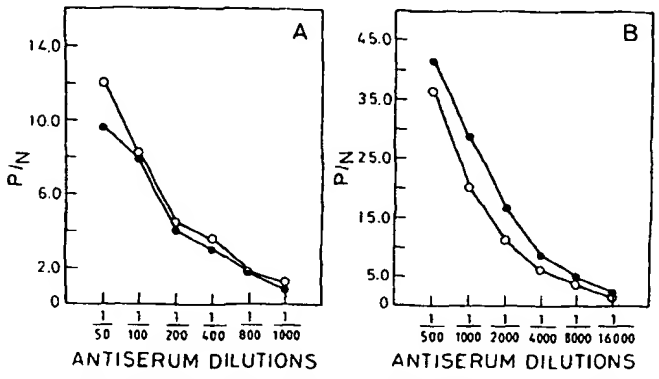

Figure 3. Titration profiles of polyclonal anti-OS[124–147] antiserum vs HBsAg and peptide. Mouse (A) or rabbit (B) anti-peptide antisera were titrated using the indicated dilutions against either peptide OS[124–147] (O) or HBsAg (●). Data are expressed as P/N ratios where N represents the absorbance obtained for reaction of a 1/100 dilution of preimmune serum with the respective Ag.

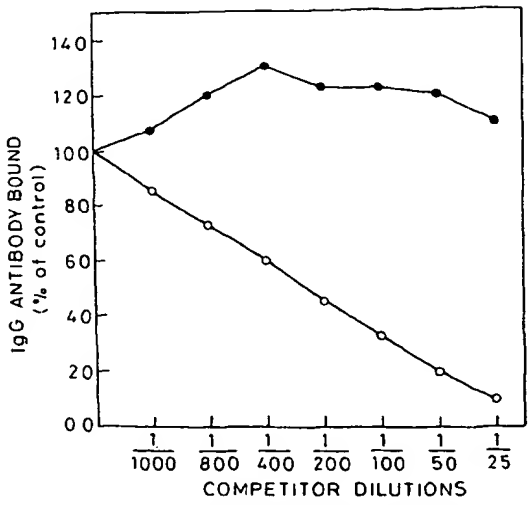

Figure 4. Competition between anti-OS[124–147] and anti-HBsAg antibodies for the OS[124–147]-related epitope on HBsAg and peptide. In one set (O) wells coated with OS[124–147] were incubated with a 1/400 final dilution of anti-HBsAg serum along with the indicated final dilutions of rabbit anti-OS[124–147] antiserum as competitor. In the second set (●) wells coated with HBsAg were incubated with a final dilution of 1/3000 of rabbit anti-OS[124–147] along with the indicated final dilutions of anti-HBsAg as competitor. Values presented are the mean of triplicate determinations and are expressed as percent of absorbance obtained in the absence of competitor.

TABLE III

Polyclonal rabbit anti-OS[124–147] antibodies recognize a variety of HBsAg subtypes[a]

HBsAg subtype	P/N	
	With anti-HBsAg	With anti-OS[124–147]
adw	10.6 ± 1.3	6.3 ± 1.7
ayw	18.2 ± 1.2	4.7 ± 1.4
ayr	5.9 ± 0.5	4.7 ± 1.5
ayw2	8.6 ± 0.2	5.4 ± 0.8
a2dw	12.2 ± 1.0	6.7 ± 1.2
adw2	9.9 ± 0.4	7.9 ± 0.1

[a] Wells coated with 1 to 2 ng of appropriate HBsAg subtype were incubated in duplicate with a 1/500 dilution of either rabbit anti-OS[124–147] or horse anti-HBsAg antisera. As negative control a 1/500 dilution of rabbit preimmune serum was used. Bound IgG was detected as described in *Materials and Methods*. Values are mean (± SD) of three independent determinations.

Anti-OS[124–147] antibodies recognize a variety of HBsAg subtypes. The data in Table I demonstrate that peptide OS[124–147] represents a conformational "a" epitope of HBsAg. To further verify this, polyclonal anti-OS[124–147] antiserum obtained in rabbits was screened for cross-reactivity with HBsAg of six different subtypes. A parallel set using polyclonal anti-HBsAg antiserum was also included as positive control. Table III shows that the anti-peptide antibodies cross-react with all the six subtypes of HBsAg tested to comparable extents, suggesting that these antibodies are primarily directed toward a conformational "a" epitope on HBsAg. This was further confirmed with a commercial kit (Hepanostika anti-HBs, Orgenon Teknika, Holland) that specifically quantitates levels of antibodies directed against the "a" determinant of HBsAg. By this procedure the rabbit anti-OS[124–147]

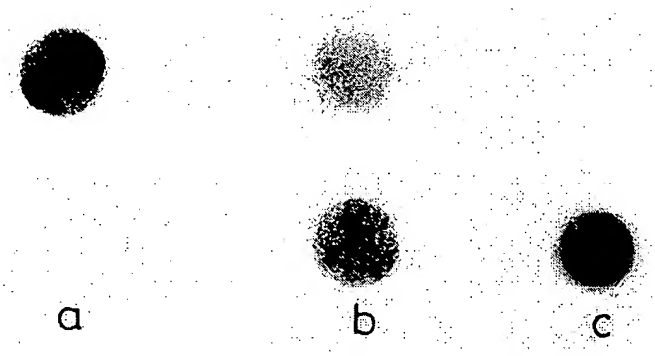

Figure 5. Anti-OS[124–147] antibodies immunoprecipitate Dane particles. Purified Dane particles were subjected to an immunoprecipitation protocol either with rabbit preimmune serum (*lane a*) or rabbit polyclonal anti-OS[124–147] serum (*lane b*) or rabbit anti-HBsAg serum (*lane c*) as described in *Materials and Methods*. HBV DNA present in the supernatant (*top*) or protein A-Sepharose pellet (*bottom*) was detected by spot hybridization.

antiserum preparation was found to contain 6500 IU/L of anti-"a" antibodies. The possible presence of a minor population of subtype specific antibodies cannot, however, be ruled out.

Anti-OS[124–147] antibodies can immunoprecipitate Dane particles. Peptide OS[124–147] elicits antibodies that recognize a conformational, group-specific determinant present on the empty envelope particles of HBsAg. It was of interest to determine whether these antibodies could also bind the same epitope on complete virions (Dane particles). This would be particularly important to ascertain from the standpoint of the potential application of OS[124–147] as a peptide vaccine for hepatitis B.

Purified Dane particles were subjected to an immunoprecipitation protocol with either rabbit anti-OS[124–147] antiserum or rabbit preimmune serum (*Materials and Methods*). Immunoprecipitation of HBV particles was detected by measuring for depletion of HBV DNA in the supernatant and its appearance in the protein A-Sepharose precipitated pellet. As shown in Figure 5, immunoprecipitation was clearly observed with anti-OS[124–147] antibodies, whereas the preimmune serum had no effect.

Thus the OS[124–147]-related epitope on the complete virion is also antigenic and the anti-peptide antibodies bind this epitope with an affinity (or avidity) that is sufficient to enable immunoprecipitation of the infectious hepatitis B virion.

DISCUSSION

Cumulative evidence described in this report demonstrates that a synthetic peptide corresponding to residues 124 to 147 of the S protein self-assembles via spontaneous oligomerization to reconstitute a conformational, group-specific antigenic determinant of HBsAg. This determinant appears to represent a dominant constituent of the epitope repertoire of HBsAg in humans. Though presently available data indicate that peptide OS[124–147] consists of a heterogenous mixture of multiple forms, a prominent band at 12 kDa was also evident. It is worthwhile to note here that the S protein from disulfide bonded HBsAg particles also displays a multiplicity of molecular weights on gel electrophoresis under nonreducing conditions (16). Full expression of HBsAg-related antigenicity was found to be stringently dependent upon the integrity of at least some of the components of the oligomerized peptide. This raises the possibility that peptide OS[124–147] may represent an epitope that is constituted by the oligomeric form of the S protein, i.e., an epitope dependent on the quarternary structure of the S protein. Alternatively the HBsAg-related epitope on peptide OS[124–147] may be localized within a single monomer sequence which adopts the right conformation only when present in the oligomerized form. Efforts are currently underway to distinguish between these two possibilities.

Attempts to map the HBsAg-related epitope on peptide OS[124–147] are ongoing in our laboratory. In recent experiments we have observed a 70 to 80% loss in HBsAg-related antigenicity of peptide OS[124–147] upon either oxidation of Met-133 or side chain methylation of Lys-141 (K. V. S. Rao and V. Manivel, unpublished results). Simultaneous chemical modification of both Met-133 and Lys-141 resulted in a greater than 90% loss in HBsAg-related antigenicity. These results are consistent with the presence of a dominant, conformational epitope that encompasses both Met-133 and Lys-141. The HBsAg-related antigenicity of peptide OS[124–147] therefore does not appear to be simply an additive composite of the previously described "a" epitopes within residues 124 to 137 (23, 24) and 139 to 147 (19). On the contrary, these results suggest the presence of a novel, conformational "a" determinant that includes amino acid residues from both regions.

Although the precise extent of closeness of "fit" (22) remains to be established, preliminary results with polyclonal anti-OS[124–147] antisera from both BALB/c mice and rabbits indicate that peptide OS[124–147] represents a fairly accurate replica of the corresponding native epitope on HBsAg. Furthermore, at least a majority of antibodies obtained in response to peptide OS[124–147] are directed against one or more conformational, disulfide-dependent epitopes. The spontaneity with which monomer S[124–147] oligomerizes to reconstitute, with high fidelity, the corresponding native epitope is particularly intriguing. An emperical analysis based on the Chou-Fasman algorithm (30, 31) indicated the presence of at least five tetrapeptide sequences (129–132, 134–137, 135–138, 142–145, and 144–147) that have a high probability of occurring in a β-turn ($p_t > 1.0 \times 10^{-4}$). It is possible that some or all of these putative turn-sequences play a critical role in folding and self-assembly of the peptide in an S protein-like manner.

Chimpanzee-challenge studies are needed to evaluate the protective nature of an anti-OS[124–147] response, but the combined immunogenicity and HBsAg-related antigenicity of OS[124–147] is indicative of its potential as a candidate component of a peptide-based vaccine for hepatitis B. In recent experiments we have observed that purified T cells from peripheral blood mononuclear cells of individuals either vaccinated with hepatitis B vaccine or acutely infected with HBV are responsive to in vitro challenge with OS[124–147], suggesting the presence of an HBsAg-relevant Th cell epitope in the peptide sequence (our unpublished results). Additionally, we have demonstrated in this report that rabbit polyclonal anti-OS[124–147] antibodies are capable of immunoprecipitating purified Dane particles. Our findings that anti-

OS[124–147] antibodies bind the corresponding native epitope with an apparent affinity higher than that of homologous antisera may have important implications in this connection.

It would be of interest to determine whether such self-assembling, conformational Ag-mimic sequences can be derived from envelope proteins of other viruses. Alternatively it may be possible to design Ag-related peptides capable of self-assembly in the manner described here. Given that a large number of viral coat proteins exist as homoaggregates in the native state (32–38), this would provide a novel and potentially useful strategy for the development of peptide vaccines against a variety of pathogens.

Acknowledgment. The authors wish to thank Ms. Radha for her excellent secretarial assistance.

REFERENCES

1. **Tiollais, P., C. Pourcel, and A. Dejean.** 1985. The hepatitis B virus. *Nature 312:489.*
2. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. *Annu. Rev. Biochem. 56:651.*
3. **Neurath, A. R., and S. B. H. Kent.** 1988. The pre-S region of hepadnavirus envelope proteins. *Adv. Virus Res. 34:64.*
4. **Neurath, A. R., S. B. H. Kent, N. Strick, P. Taylor, and C. E. Stevens.** 1985. Hepatitis B virus contains pre-S gene-encoded domains. *Nature 315:154.*
5. **Rao, K. R., and G. N. Vyas.** 1973. Hepatitis B antigen activity in protein subunits produced by sonication. *Nature 241:240.*
6. **Melnick, J. L., G. R. Dreesman, and F. B. Hollinger.** 1976. Approaching the control of viral hepatitis type B. *J. Infect. Dis. 133:210.*
7. **Tiollais, P., P. Charnay, and G. N. Vyas.** 1981. Biology of hepatitis B virus. *Science 213:403.*
8. **Neurath, A. R., S. B. H. Kent, N. Strick, and K. Parker.** 1986. Identification and Chemical synthesis of a host cell receptor binding site on hepatitis B virus. *Cell 46:429.*
9. **Neurath, A. R., S. B. H. Kent, N. Strick, and K. Parker.** 1988. Vaccination with synthetic hepatitis B virus peptides. *Appl. Virol. Res. 1:107.*
10. **Neurath, A. R., B. Seto, and N. Strick.** 1989. Antibodies to synthetic peptides from the preS1 region of the hepatitis B virus (HBV) envelope (env) protein are virus-neutralizing and protective. *Vaccine 7:234.*
11. **Neurath, A. R., S. B. H. Kent, K. Parker, A. M. Prince, N. Strick, B. Brotman, and P. Sprout.** 1986. Antibodies to a synthetic peptide from the pre S 120–145 region of the hepatitis B virus envelope are virus neutralizing. *Vaccine 4:35.*
12. **Itoh, Y., E. Takai, H. Ohnuma, K. Kitajima, F. Tsuda, A. Machida, S. Mishiro, T. Nakamura, Y. Miyakawa, and M. Mayumi.** 1986. *Proc. Natl. Acad. Sci. USA 83:9174.*
13. **Sukeno, N., R. Shirachi, J. Yamaguchi, and N. Ishida.** 1972. Reduction and reoxidation of Australia antigen: loss and reconstitution of particulate structure and antigenicity. *J. Virol. 9:182.*
14. **Vyas, G. N., K. R. Rao, and A. B. Ibrahim.** 1972. Australia antigen (hepatitis B surface antigen): a conformational antigen dependent on disulfide bonds. *Science 178:1300.*
15. **Imai, M., A. Gotob, K. Nishioka, S. Kurashima, Y. Miyakawa, and M. Mayumi.** 1974. Antigenicity of reduced and alkylated Australia antigen. *J. Immunol. 112:416.*
16. **Wampler, D. E., E. D. Lehman, J. Boger, W. J. McAleer, and E. J. Scolnick.** 1985. Multiple chemical forms of hepatitis B surface antigen produced in yeast. *Proc. Natl. Acad. Sci. USA 82:6830.*
17. **LeBouvier, G. L., R. W. McCollum, W. J. Hierholzer, G. R. Irwin, S. Krugman, and J. P. Giles.** 1972. Subtypes of Australia antigen and hepatitis-B virus. *JAMA 222:928.*
18. **Gerin, J. L., H. Alexander, J. W. Shih, R. H. Purcell, G. Dapolito, R. Engle, N. Greene, J. G. Sutcliffe, T. M. Shinnick, and R. A. Lerner.** 1983. Chemically synthesized peptides of hepatitis B surface antigen duplicate the d/y specificities and induce subtype-specific antibodies in chimpanzees. *Proc. Natl. Acad. Sci. USA 80:2365.*
19. **Bhatnagar, P. K., E. Papas, H. E. Blum, D. R. Milich, D. Nitaki, M. J. Karels, and G. N. Vyas.** 1982. Immune response to synthetic peptide analogues of hepatitis B surface antigen specific for the a determinant. *Proc. Natl. Acad. Sci. USA 79:4400.*
20. **Prince, A. M., H. Ikram, and T. P. Hopp.** 1982. Hepatitis B virus vaccine: identification of HBsAg/a and HBsAg/d but not HBsAg/y subtype antigenic determinants on synthetic immunogenic peptide. *Proc. Natl. Acad. Sci. USA 79:579.*
21. **Brown, S. E., C. R. Howard, A. J. Zuckerman, and M. W. Steward.** 1984. Determinsation of affinity of antibodies to hepatitis B surface antigen in human sera. *J. Immunol. Methods 72:41.*
22. **Howard, C., S. Brown, B. Sisley, C. Stanley, S-H. Chen, and M. Steward.** 1987. Cellular and antibody responses to HBV S and pre-S peptides in man. In *Hepadna Viruses.* W. Robinson, K. Kolke and H. Will, eds. Alan R. Liss, New York, p. 495.
23. **Dreesman, G. R., Y. Sanchez, I. Ionescu-Matiu, J. T. Sparrow, H. R. Six, D. L. Peterson, F. B. Hollinzer, and J. L. Melnick.** 1982. Antibody to hepatitis B surface antigen after a single inoculation of uncoupled synthetic hepatitis B surface antigen peptides. *Nature 295:158.*
24. **Ionescu-Matiu, I. R. C. Kennedy, J. T. Sparrow, A. R. Culwell, Y. Sanchez, J. L. Melnick, and G. R. Dreesman.** 1983. Epitopes associated with a synthetic hepatitis B surface antigen peptide. *J. Immunol. 130:1947.*
25. **Howard, C., H. Stirk, A. Buckley, S. Brown, and M. Steward.** 1989. The use of synthetic peptides for the analysis of the envelope of the hepatitis B virus (HBV). In *Synthetic Peptides: Approaches to Biological Problems.* J. P. Tam and E. T. Kaiser, eds. Alan R. Liss, New York, p. 211.
26. **Merrifield, R. B.** 1963. Peptide Synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc. 85:2149.*
27. **Landers, T. A., H. B. Greenberg, and W. S. Robinson.** 1977. Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. *J. Virol. 23:368.*
28. **Scotts, J., M. Hadchouel, C. Hery, J. Yvart, P. Tiollais, and C. Brachol.** 1983. Detection of hepatitis B virus DNA in the serum by a simple spot hybridisation technique:Comparison with results for other viral markers. *Hepatology 3:279.*
29. **Ellman, G. L.** 1959. Tissue sulfhydryl groups. *Arch. Biochem. Biophys. 82:70.*
30. **Chou, P. Y., and G. D. Fasman.** 1978. Experimental predictions of protein conformation. *Annu. Rev. Biochem. 47:251.*
31. **Chou, P. Y., and G. D. Fasman.** 1979. Prediction of B-turns. *Biophys. J. 26:367.*
32. **Pinter, A., W. J. Honnen, S. A. Tilley, C. Bona, H. Zaghouani, M. K. Gorny, and S. Zolla-Pazner.** 1989. Oligomeric structure of gp 41 the transmembrane protein of human immunodeficiency virus type 1. *J. Virol. 63:2674.*
33. **Pinter, A., and E. Fleissner.** 1979. Characterization of oligomeric complexes of murine and feline leukemia virus envelope and core components formed upon crosslinking. *J. Virol. 30:157.*
34. **Racevskis, J., and N. Sarkar.** 1980. Murine mammary tumor virus structural protein interactions: formation of oligomeric complexes with cleavable cross-linking agents. *J. Virol. 35:833.*
35. **Doms, R. W., and A. Helenius.** 1986. Quarternary structure of influenza virus hemagglutinin after acid treatment. *J. Virol. 60:833.*
36. **Einfield, D., and E. Hunter.** 1988. Oligomeric structure of a prototype retrovirus glycoprotein. *Proc. Natl. Acad. Sci. USA 85:8688.*
37. **Varghese, J. N., W. G. Laver, and P. M. Colman.** 1983. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution. *Nature 303:35.*
38. **Kries, T. E., and H. F. Lodish.** 1986. Oligomerization is essential for transport of vesicular stomatis viral glycoproteins to the cell surface. *Cell 46:929.*

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 6509–6513, July 1995
Biochemistry

Incorporation of glutamine repeats makes protein oligomerize: Implications for neurodegenerative diseases

KELVIN STOTT*, JONATHAN M. BLACKBURN*, P. J. G. BUTLER†, AND MAX PERUTZ*†

*Medical Research Council Centre for Protein Engineering and †Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Contributed by Max Perutz, April 11, 1995

ABSTRACT **Many transcription factors and some other proteins contain glutamine repeats; their abnormal expansion has been linked to several dominantly inherited neurodegenerative diseases. Having found that poly(L-glutamine) alone forms β-strands held together by hydrogen bonds between their amide groups, we surmised that glutamine repeats may form polar zippers, an unusual motif for protein–protein interactions. To test this hypothesis, we have engineered a Gly-Gln_{10}-Gly peptide into the inhibitory loop of truncated chymotrypsin inhibitor 2 (CI2), a small protein from barley seeds, by both insertion and replacement. Gel filtration resolved both mutant inhibitors into at least three fractions, which analytical ultracentrifugation identified as monomers, dimers, and trimers of the recombinant protein; the truncated wild-type CI2 formed only monomers. CD difference spectra of the dimers and trimers versus wild type indicated that their glutamine repeats formed β-pleated sheets, while those of the monomers versus wild type were more suggestive of type I β-turns. The CD spectra of all three fractions remained unchanged even after incubation at 70°C; neither the dimers nor the trimers dissociated at this temperature. We argue that the stability of all three fractions is due to the multiplicity of hydrogen bonds between extended strands of glutamine repeats in the oligomers or within a β-hairpin formed by the single glutamine repeat of each monomer. Pathological effects may arise when expanded glutamine repeats cause proteins to acquire excessively high affinities for each other or for other proteins with glutamine repeats.**

Five dominantly inherited neurodegenerative diseases have now been linked to abnormally expanded stretches of polyglutamine in the affected proteins: Huntington disease; spinal and bulbar muscular atrophy, also known as Kennedy disease; spinocerebellar ataxia type 1; dentatorubral–pallidoluysian atrophy (1, 2); and Machado–Joseph disease (3). All five diseases become more severe and begin earlier the longer the glutamine repeats. They are encoded by CAG trinucleotide repeats, which tend to lengthen in successive generations of affected individuals, especially in male transmission. After modeling had suggested that glutamine repeats might act as polar zippers, linking β-strands by hydrogen bonds between both their main-chain and side-chain amides, a synthetic peptide with the sequence Asp_2-Gln_{15}-Lys_2 was found to aggregate into pleated β-sheets (4–6). Encouraged by this result, we have now engineered glutamine repeats into the inhibitory loop of chymotrypsin inhibitor 2 (CI2) from barley seeds, a small monomeric protein of known structure (7), in order to determine whether they make this protein associate into oligomers.

We chose this small protein rather than any of those affected by the five neurodegenerative diseases, because none of those has yet been isolated in workable quantities. They are all very large, and it seemed doubtful that they could be expressed in soluble and native form by recombinant DNA technology. At this early stage, it seemed important above all to establish in principle whether incorporation of glutamine repeats makes proteins oligomerize, for which there has previously been no direct evidence from protein chemistry. Glutamine-rich domains have been found to be responsible for interaction between transcription factors, but it has not been proved that this is due solely to specific linkages between their glutamines.

MATERIALS AND METHODS

pJB2 was constructed by ligation of the 200-bp *Nde* I/*Hin*dIII fragment of pCI2 (8) with the 4.5-kb *Nde* I/*Hin*dIII fragment of pRH1090 (9). pCI2-Q10(ins), the vector used for expression of the loop insertion mutant of truncated CI2, was constructed by insertion of a synthetic DNA fragment encoding the Gly-Gln_{10}-Gly peptide into the *Nco* I restriction site at the Met-59 codon of the CI2 gene in pJB2; the insert was first generated by phosphorylating and annealing the two 39-mer oligonucleotides 5′-CATGGGTCAGCAGCAACAACAG-CAGCAACAACAGCAGGG-3′ and 5′-CATGCCCTGCT-GTTGTTGCTGCTGTTGTTGCTGCTGACC-3′ and was then ligated directly into a dephosphorylated, gel-purified partial *Nco* I digest of pJB2. pCI2-Q10(rep), the vector used for expression of the loop replacement mutant of truncated CI2, was constructed by substitution of the *Pst* I/*Hin*dIII fragment of the CI2 gene in pJB2 with an analogous fragment in which codons 54–61 were replaced with a stretch of DNA encoding the Gly-Gln_{10}-Gly peptide; the 140-bp substitute fragment was first generated by PCR, using pJB2 as template DNA and the two oligonucleotides 5′-AAAACTGCAGGA-CAAGCCAGAGGCGCAAATCATAGTTCTGCCGGTG-GGGCAGCAGCAACAACAGCAGCAACAACAGCAGG-GGCGGATCGACCGCGTCCG-3′ and 5′-GCCGCCAGGC-AAATTCTG-3′ as forward and reverse primers, respectively; after complete digestion with *Pst* I and *Hin*dIII, the purified fragment was ligated with the 4.6-kb *Pst* I/*Hin*dIII fragment of pJB2. For each of the three constructs, ligated DNA was electroporated into *Escherichia coli* strain NM554, cells were plated on rich medium containing chloramphenicol, colonies were screened by PCR, and one positive clone was restriction-mapped and sequenced to confirm its identity (10).

The soluble extracts, each from 1 liter of induced culture, were first precipitated with 70% saturated ammonium sulfate, dialyzed into 50 mM Tris·HCl (pH 8.6), passed through 1 ml of DE52 anion-exchange resin, and then concentrated to ≈5 ml over ultrafiltration membranes with a molecular weight cutoff of 3000. NaCl (150 mM) was added to the concentrated extracts before loading onto a HiLoad 26/60 Superdex 75 size-exclusion column for gel filtration by fast protein liquid chromatography. Protein was eluted with 50 mM Tris·HCl, pH 8.6/150 mM NaCl at 2.5 ml/min and monitored as absorbance at 280 nm.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: CI2, chymotrypsin inhibitor 2.

RESULTS

CI2 consists of a single polypeptide chain of 83 residues, the first 18 of which are disordered in its crystal structure (Fig. 1); they are followed by a short β-strand, a three-turn α-helix, and another three β-strands associated into a pleated sheet, which forms a stable hydrophobic core with the α-helix. β-strands 2 and 3 are connected by a long and flexible solvent-exposed loop consisting of residues 54–62, including the inhibitory methionine (Met-59). We have made two mutants of a truncated form of CI2 from which the first 20 residues were deleted. In the first mutant, a sequence of 10 glutamine residues flanked by two glycines (for flexibility) was inserted into the loop at Met-59, while in the second mutant, residues 54–61 of the loop were replaced with this sequence. The parent vector pJB2 (Fig. 2) was constructed for expression of the truncated wild-type CI2 and was then used to construct vectors for expression of the two mutant inhibitors. Expression was induced in *E. coli* strain NM554 with 1 mM isopropyl β-D-thiogalactopyranoside, yielding ≈150 mg of recombinant protein per liter of culture after purification.

Fig. 3 shows the fractionation of soluble cell extracts of the truncated wild-type inhibitor and of the two mutants by gel filtration. Cell extract of the truncated wild-type inhibitor gave a broad heterogeneous fraction of large cellular proteins (fraction a), a clean concentrated fraction of the recombinant protein (fraction d), and a series of low molecular weight fractions (fraction e), which were later found to be nucleic acid, probably RNA. The cell extracts of the two mutant inhibitors, on the other hand, both yielded two additional fractions (fractions b and c) of higher apparent molecular weight than the main fraction of monomeric recombinant protein, but quite distinct from the broad heterogeneous fraction of large

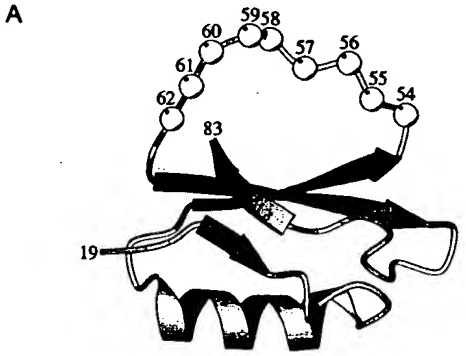

FIG. 1. (*A*) Structure of wild-type CI2; the first 18 residues are disordered and are not shown (7). Loop residues 54–62 are represented by spheres. The Gly-Gln$_{10}$-Gly peptide was inserted into the loop at Met-59 in one mutant, while the second mutant was generated by replacing residues 54–61 of the loop with this peptide. (*B*) Amino acid sequence of the truncated wild-type CI2 and of the two mutants. Original residues of the loop are underlined.

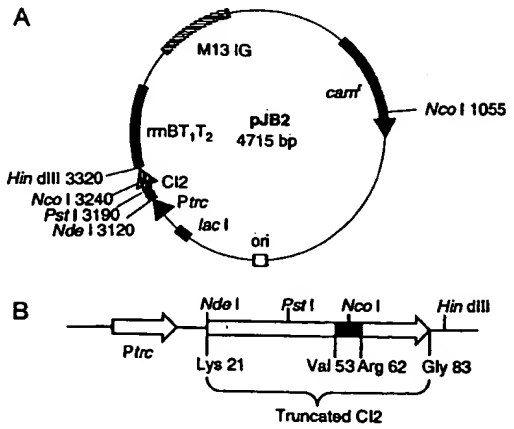

FIG. 2. (*A*) Plasmid map of the parent vector pJB2 used for expression of the truncated wild-type CI2. (*B*) Structure of the truncated CI2 gene in pJB2, including the relevant codons and restriction sites.

cellular proteins. SDS/PAGE of fractions b, c, and d of each mutant inhibitor produced a single clean band at the same position, close to that of the truncated wild-type inhibitor, showing that all three fractions consisted of the recombinant protein. This was confirmed by electrospray mass spectrometry, which gave essentially identical values for the molecular weights of each of the three fractions of the two mutant inhibitors, indicating that fractions b and c were higher-order aggregates of the recombinant proteins in fraction d. The observed molecular weights also verified their correct expression (Table 1).

The order of oligomerization in fractions b, c, and d of both mutant inhibitors was determined by analytical ultracentrifugation. Fig. 4 shows the results for the loop insertion mutant inhibitor, indicating apparent molecular weights for fractions b, c, and d of 26,000, 17,900, and 9900, respectively. These weights agree with those expected for trimer, dimer, and

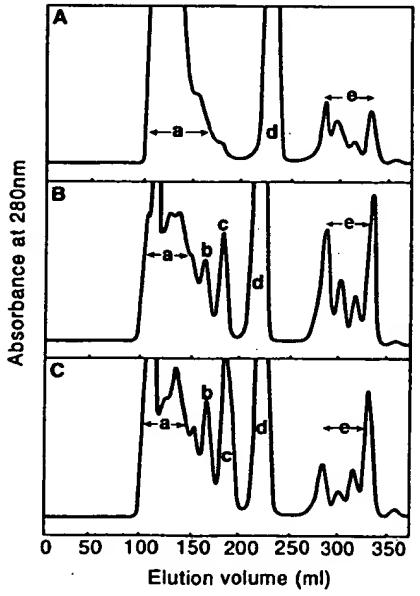

FIG. 3. Preparative gel filtration of soluble cell extracts from *E. coli* expressing the truncated wild-type CI2 (*A*), the loop insertion mutant (*B*), and the loop replacement mutant (*C*). Full absorbance scale of each elution profile is 2.0.

Table 1. Calculated vs. observed molecular weight

	Molecular weight	
	Calculated	Observed
Truncated wild-type CI2	7303.6	7304.1 (±0.7)
Loop insertion mutant	8830.3	8830.3 (±0.5)
Loop replacement mutant	7804.0	7804.0 (±0.3)

monomer (M_r, 26,500, 17,700, and 8,800). Similar results were obtained for the loop replacement mutant inhibitor (data not shown). Analysis of the fractions by fast protein liquid chromatography (Pharmacia) on a calibrated Superdex 75 HR 10/30 gel-filtration column produced peaks at elution volumes that were also compatible with the apparent molecular weights expected for trimer, dimer, and monomer (data not shown). Fractions c and d of each mutant inhibitor produced clean single peaks, whereas fraction b produced a peak with a small shoulder coinciding with the peak of fraction c, probably due to dimer as a contaminant from the less efficient isolation of the trimeric fraction. This is reflected by the decrease in apparent molecular weight of fraction b from trimer to dimer at low protein concentration in Fig. 4.

The CD spectra of all three fractions of either mutant were similar to the spectrum of the truncated wild-type inhibitor, as shown in Fig. 5 for the loop insertion mutant, which speaks against any major changes in tertiary structure. The difference spectrum between the trimer and the wild type shows a positive peak at 195 nm and a negative peak at 223 nm, coincident with the pleated β-sheet spectrum of the original oligopeptide Asp_2-Gln_{15}-Lys_2 (5). The difference spectrum of monomer versus wild type above 200 nm resembles the CD spectrum of type I β-turns, but it lacks its large positive peak at 195 nm (11). We have been able to reproduce it by heating a solution of the oligopeptide to 95°C and letting it cool slowly to 25°C (Fig. 5*C*). Heating may have caused the β-strands in the pleated sheets of the oligopeptide to dissociate and fold into hairpins; gel filtration revealed that these then reassociated on cooling. The similarity between the spectrum of the heated oligopeptide and the difference spectrum of the monomer versus wild type suggests that the glutamine repeat inserted in the CI2 may also fold into a hairpin. The difference spectrum of the dimer

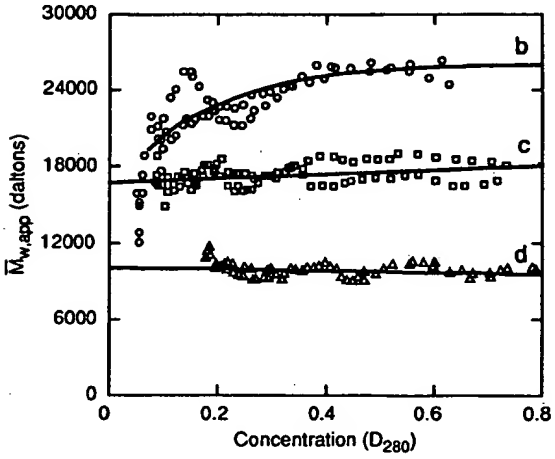

FIG. 4. Plots of apparent weight average molecular weights ($\bar{M}_{w,app}$) against concentration for the peaks b, c, and d from fast protein liquid chromatography fractionation. Molecular weights were determined by equilibrium centrifugation and analysis as described (12), assuming a partial specific volume of 0.73 ml/g. Curve b was fitted for nonideality, while curves c and d were drawn in arbitrarily to fit the overlaid data from separate experiments.

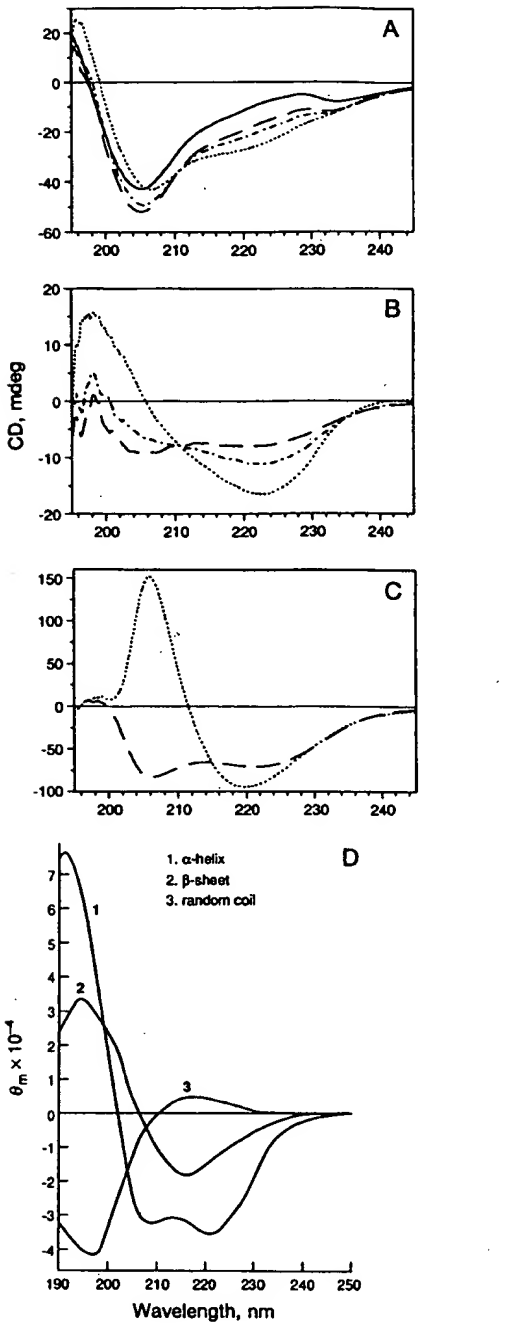

FIG. 5. (*A*) CD spectra of truncated wild-type CI2 (——) and of monomeric (— —), dimeric (- - -), and trimeric (- - -) fractions of the loop insertion mutant in 50 mM potassium phosphate buffer (pH 6.85) at 25°C. All three fractions and the wild type had an absorbance of 0.4 at 282 nm. With an extinction coefficient calculated from their amino acid composition of 6970 $M^{-1}cm^{-1}$, which was independent of the presence of the glutamines and of 6 M guanidinium chloride, this absorbance corresponded to a concentration of 57.4 μM. (*B*) Corresponding difference spectra of the three mutant fractions relative to the spectrum of the truncated wild-type CI2. (*C*) Spectrum of the oligopeptide Asp_2-Gln_{15}-Lys_2 before (- - -) and after (——) heating to 95°C. Spectra were recorded as the average of 30 individual spectra, each measured using a Jasco J-720 spectropolarimeter with a 0.1-cm-pathlength cell, at a scan rate of 20 nm/min. (*D*) Standard CD spectra of α-helix, β-sheet, and random coil structures of poly(L-lysine). (Reproduced with permission from ref. 13; copyright 1969, American Chemical Society.)

versus wild type is intermediate between the difference spectra of the trimer and the monomer versus wild type.

That the three different fractions of either mutant could be isolated suggests that they are all fairly stable; otherwise, a single fraction containing an equilibrium mixture of the three oligomerization states would have been produced. In fact, analytical gel filtration revealed that all three fractions of the loop insertion mutant remained pure even after storage at 4°C for 2 months. Those of the loop replacement mutant inhibitor were slightly less stable; over a period of weeks, a small fraction of the trimers of this mutant dissociated into monomers and dimers, and a small fraction of the dimers dissociated into monomers, while the monomers remained pure, even at a concentration of 2 mM. To compare the stability of the three fractions with that of the wild type, we followed their CD at 222 nm while raising the temperature from 2°C to 95°C (Fig. 6). The wild type was stable up to 80°C and denatured at 85°C, when an increase in absorption at 280 nm showed that it aggregated. The CD of the three fractions exhibited no transitions below 50°C, which showed that their glutamine zippers were stable to at least this temperature. The monomer showed only a single transition at 80°C, when it denatured and aggregated, as did the dimer and trimer, such that the CD curves of all three fractions converged at 90°C. The CD curves of the dimer and trimer indicate that they underwent structural transitions between 50°C and 70°C. To determine whether these corresponded to dissociation into monomers, we incubated all three fractions at 70°C for 1 hr and then cooled them slowly to 25°C, but instead of finding the CD spectra of the dimer and trimer converted to the spectrum of the monomer, they all remained unchanged, proving that their glutamine zippers remained intact even at 70°C. When we tried to establish a dynamic equilibrium between monomers and oligomers in various concentrations of urea, we found that the entire protein denatured cooperatively with its glutamine repeat at around 3.0 M urea.

The stability of all three fractions implies that both association and dissociation rates are extremely slow. The slow association rates of the monomers can be explained only by the presence of intramolecular hydrogen bonds within their glutamine repeats, which are likely to be bent into hairpins. The slow dissociation rates of the oligomers are clearly due to the multiplicity of intermolecular hydrogen bonds between extended β-strands of glutamine repeats. The conformation of β-strands would be the same in hairpins as in β-sheets, but hairpins in the monomers would be entropically favored over β-sheets of glutamine repeats in the oligomers. This may explain why monomers were the dominant species formed during overexpression in *E. coli*. In low expression, oligomers would be less likely to form, unless the glutamine repeats become long enough for the hairpins themselves to associate into β-sheets.

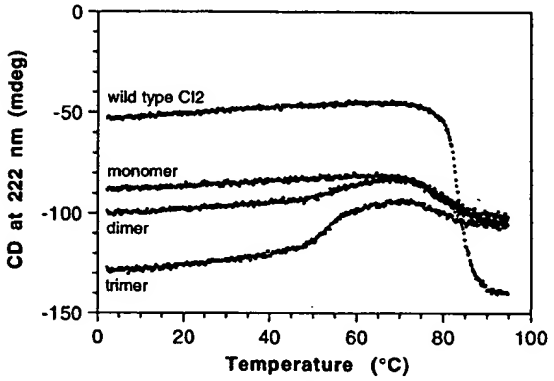

FIG. 6. Thermal denaturation of truncated wild-type CI2 and of the three fractions of the loop insertion mutant, each at a concentration of 28.7 μM in 50 mM potassium phosphate buffer at pH 6.85, followed by CD at 222 nm using a cell of 1-cm pathlength.

DISCUSSION

How common are glutamine repeats or glutamine-rich domains, where do they occur, and what is their function? On screening the Swiss-Prot data base for proteins with at least 20 glutamine repeats, Gerber *et al.* (14) found 33 of 40 top scoring proteins to be transcription factors. Tjian and his colleagues (15–19) have explored the role of glutamine-rich domains in the human transcription factor SP1. Transcription of reporter genes was enhanced by interaction between SP1s bound to GC-rich promoters 1.8 kb apart; the glutamine-rich domains proved essential for this interaction, which was vividly demonstrated in electron micrographs picturing the looping of the intervening DNA when two distantly bound SP1s were joined together. SP1 also interacted strongly with glutamine-rich TATA-binding-associated factors. Gerber *et al.* (14) measured activation of transcription by the Gal-4 chimeric protein, one part of which was linked to SP1 and the other to TATA-binding-associated factors or other proteins with either glutamine repeats or glutamine-rich domains. They all stimulated high levels of transcription. Stimulation increased with the number of glutamine repeats up to 40, remained constant up to 80 repeats, and then declined. These experiments suggest that glutamine repeats play a role in transcription similar to that of leucine zippers.

Only one of the proteins responsible for the five dominantly inherited neurodegenerative diseases is known to be a transcription factor. This is the androgen receptor responsible for Kennedy disease. The function of the Huntington disease protein is still unknown, because its amino acid sequence of >3100 residues (M_r, ≈350,000) shows no homology with other known proteins; contrary to earlier reports, it has now been located in the cytoplasm rather than the cell nucleus (20). C. A. Ross (Johns Hopkins Department of Psychiatry and Neuroscience) has informed us that a Western blot of monkey cortex fractionated on a nondenaturing gel and probed with anti-Huntington disease antibodies by G. Schilling and A. H. Sharp yielded a broad peak at M_r ≈700,000, showing that the protein either associated into dimers or associated with other proteins.

Several observations have proved that neural damage in diseases affected by expanded glutamine repeats is due to gain rather than loss of function. Our results show that glutamine repeats make proteins associate into stable oligomers. In transcription factors, expanded repeats may therefore lead to "aberrant transcriptional activity" (21), because these repeats acquire excessive affinities for each other or for complementary regulatory proteins with glutamine repeats. In cytoplasmic proteins, expanded glutamine repeats may lead to "wrong" interactions with other proteins. We hope that our proof of the function of glutamine repeats will serve as a step toward unraveling the molecular mechanism of these terribly distressing neurodegenerative diseases. Our discovery may also find uses in protein engineering; oligomers held together by glutamine repeats may serve as models for the design of new proteins.

We thank Prof. A. R. Fersht for suggesting the use of CI2 for this experiment, Fiona Suit for synthesizing our oligonucleotide primers, Dr. Ian Fearnley for determining the molecular weights of our constructs by electrospray mass spectrometry, Dr. Olga Perisic for helpful advice with their preparation, and Dr. Tony Johnson for determining the distribution of molecular weights of our oligopeptide before and after heating by gel filtration. K.S. is supported by a Medical Research Council Research Studentship, J.M.B. is supported by a Fitzwilliam College Research Fellowship, and M.P.'s research is supported by National Institutes of Health Grant HL31461.

1. Martin, J. B. (1993) *Science* **262**, 674–676.
2. Ross, C. A., McInnes, M. G., Margolis, R. L. & Li, S.-H. (1993) *Trends Neurosci.* **16**, 254–259.
3. Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I., Kimura, J., Narumiya, S. & Kakizuka, A. (1994) *Nat. Genet.* **8**, 221–228.
4. Perutz, M. P., Staden, R., Moens, L. & De Baere, I. (1993) *Curr. Biol.* **3**, 249–253.
5. Perutz, M. F., Johnson, T., Suzuki, M. & Finch, J. T. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5355–5358.
6. Perutz, M. F. (1994) *Protein Sci.* **3**, 1629–1637.
7. McPhalen, C. A., Svendsen, I., Jonassen, I. & James, M. N. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7242–7246.
8. Jackson, S. E., Moracci, M., el Masri, N., Johnson, C. M. & Fersht, A. R. (1993) *Biochemistry* **32**, 11259–11269.
9. Baldwin, J. E., Blackburn, J. M., Heath, R. I. & Sutherland, J. D. (1992) *Bioorg. Med. Chem. Lett.* **2**, 663–668.
10. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
11. Perczel, A., Hollósi, M., Sündor, P. & Fasman, G. D. (1993) *Int. J. Pept. Protein Res.* **41**, 223–236.
12. Tennent, G. A., Butler, P. J. G., Hutton, T., Woolfitt, A. R., Harvey, D. J., Rademacher, T. W. & Pepys, M. B. (1993) *Eur. J. Biochem.* **214**, 91–97.
13. Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* **8**, 4108–4116.
14. Gerber, H. P., Seipel, K., Georgiev, O., Höfferer, M., Hug, M., Rusconi, S. & Schaffner, W. (1994) *Science* **263**, 808–811.
15. Courey, A. J. & Tjian, R. (1988) *Cell* **55**, 887–898.
16. Courey, A. J., Holtzman, D. A., Jackson, S. P. & Tjian, R. (1989) *Cell* **59**, 827–836.
17. Pascal, E. & Tjian, R. (1991) *Genes Dev.* **5**, 1646–1656.
18. Su, W., Jackson, S., Tjian, R. & Echols, H. (1991) *Genes Dev.* **5**, 820–826.
19. Hoey, T. H., Weinzierl, R. O. J., Gill, G., Chen, J.-L., Dynlacht, B. D. & Tjian, R. (1993) *Cell* **72**, 247–260.
20. Sharp, A. H., Loev, S. J., Schilling, G., Steiner, J. P., Lo, A., Hedreen, J., Sisodia, S., Dawson, T. M., Snyder, S. H. & Ross, C. A. (1994) *Soc. Neurosci. Abstr.* **20**, 1648.
21. Mhatre, A. N., Trifiro, M. A., Kaufman, M., Kazemi-Esfarani, P., Figlewicz, D., Rouleau, G. & Pinsky, L. (1993) *Nat. Genet.* **5**, 184–188.

Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 6953–6957, October 1987
Neurobiology

Synthetic peptide homologous to β protein from Alzheimer disease forms amyloid-like fibrils *in vitro*

(aging/fibrous proteins/protein conformation/amyloidosis/x-ray diffraction)

DANIEL A. KIRSCHNER*†‡, HIDEYO INOUYE*, LAWRENCE K. DUFFY†§, ALISON SINCLAIR*, MARCIA LIND*, AND DENNIS J. SELKOE†§

*Department of Neuroscience, Children's Hospital, Boston, MA 02115; §Center for Neurologic Diseases, Brigham and Women's Hospital, Boston, MA 02115; and †Department of Neuropathology and Neurology, Harvard Medical School, Boston, MA 02115

Communicated by Richard L. Sidman, June 18, 1987 (received for review April 1, 1987)

ABSTRACT **Progressive amyloid deposition in senile plaques and cortical blood vessels may play a central role in the pathogenesis of Alzheimer disease. We have used x-ray diffraction and electron microscopy to study the molecular organization and morphology of macromolecular assemblies formed by three synthetic peptides homologous to β protein of brain amyloid: β-(1–28), residues 1–28 of the β protein; [Ala^{16}]β-(1–28), β-(1–28) with alanine substituted for lysine at position 16; and β-(18–28), residues 18–28 of the β protein. β-(1–28) readily formed fibrils *in vitro* that were similar in ultrastructure to the *in vivo* amyloid and aggregated into large bundles resembling those of senile plaque cores. X-ray patterns from partially dried, oriented pellets showed a cross-β-conformation. A series of small-angle, equatorial maxima were consistent with a tubular fibril having a mean diameter of 86 Å and a wall composed of pairs of cross-β-pleated sheets. The data may also be consistent with pairs of cross-β-sheets that are centered 71-Å apart. [Ala^{16}]β-(1–28) formed β-pleated sheet assemblies that were dissimilar to *in vivo* fibrils. The width of the 10-Å spacing indicated stacks of about six sheets. Thus, substitution of the uncharged alanine for the positively charged lysine in the β-strand region enhances the packing of the sheets and dramatically alters the type of macromolecular aggregate formed. β-(18–28) formed assemblies that had even a greater number of stacked sheets, ≈24 per diffracting domain as indicated by the sharp intersheet reflection. Our findings on these homologous synthetic assemblies help to define the specific sequence that is required to form Alzheimer-type amyloid fibrils, thus providing an *in vitro* model of age-related cerebral amyloidogenesis.**

The major histopathological features of Alzheimer disease (AD) are neurofibrillary tangles, neuritic or senile plaques, and amyloid angiopathy (1). The plaques contain cores of extracellular proteinaceous filaments that have been identified as amyloid, based on their green birefringence after staining with Congo red and their 40- to 90-Å diameter (2, 3). Structurally similar amyloid filaments also occur in the walls of some capillaries, arterioles, and small arteries in the cerebral cortex and in some meningeal arteries in patients with AD (4–10). These amyloid deposits are also found in the brains of Down syndrome patients over 30 years of age (11), as well as in the brains of aged normal humans (12) and lower mammals (13).

The major proteinaceous component of amyloid both from the cerebral blood vessel walls and the plaque cores is an ≈4-kDa protein (8, 9, 14–17) designated the β protein (15) or A4 (8). The first 28 amino acids of this protein have been sequenced (7). A gene coding for the β-amyloid protein has been mapped to chromosome 21 of the human genome (18–20), and the complete sequence of the putative precursor protein has been determined (21).

The current study was carried out to investigate the amyloidogenic properties of peptides sharing sequences with the β protein. Electron microscopy was used to examine the morphology of the peptide assemblies, and x-ray diffraction was used to examine their molecular organization. We have found that a synthetic polypeptide identical in sequence to the N-terminal 28 residues of the β protein readily forms fibrils *in vitro*. Electron microscopy reveals that these fibrils are ultrastructurally similar to the *in vivo* amyloid fibrils and aggregate into bundles resembling amyloid deposits in plaque cores and vessels. X-ray diffraction patterns from unfixed, oriented pellets of the reassembled fibrils show details of structural organization not previously observed, to the best of our knowledge, in patterns from either cerebral or systemic amyloids. Analysis of the patterns suggests that the fibril either is a hollow, tubular structure 86 Å in diameter or consists of slab-like walls centered 71 Å apart. In either case, the wall of the fibril is composed of pairs of β-sheets with the peptide chains running approximately perpendicular to the fibril axis. Since chemical modification of the lysines of this peptide reduces its ability to inhibit antibody binding to amyloid (22), we also examined a 28-residue peptide in which an alanine was substituted for lysine-16. This peptide formed linear assemblies that were distinctly different from those formed by the β protein and that showed much less propensity to associate in bundles. Finally, a synthetic peptide composed of residues 18–28 formed assemblies having extensive intersheet stacking.

METHODS

Synthesis and Purification of Peptides. Polypeptides were synthesized using an ABI Synthesizer model 380B (Applied Biosystems, Foster City, CA) and purified by reverse-phase HPLC using a C_4 column (Vydac, Hesperia, CA). Their amino acid sequences were confirmed by sequencing. In this paper, we use β-(1–28) to denote the synthetic peptide consisting of residues 1–28 of the AD amyloid β protein, [Ala^{16}]β-(1–28) to denote the synthetic peptide identical to β-(1–28), except that alanine has been substituted for lysine-16, and β-(18–28) to denote the peptide with residues 18–28 from the β protein. The β-(1–28) peptide solution at ≈2 mg/ml, aqueous concentration, precipitated during storage at 4°C forming a gel with a concentration of ≈20 mg/ml. Aliquots of this gel were used for thin-section electron microscopy and x-ray diffraction and were diluted for negative staining. The [Ala^{16}]β-(1–28) solution, which also had an initial protein concentration of ≈2 mg/ml, formed a small amount of fine precipitate that was examined in thin sections and by x-ray

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: AD, Alzheimer disease.

‡To whom reprint requests should be addressed at: Department of Neuroscience, Children's Hospital, 300 Longwood Avenue, Boston, MA 02115.

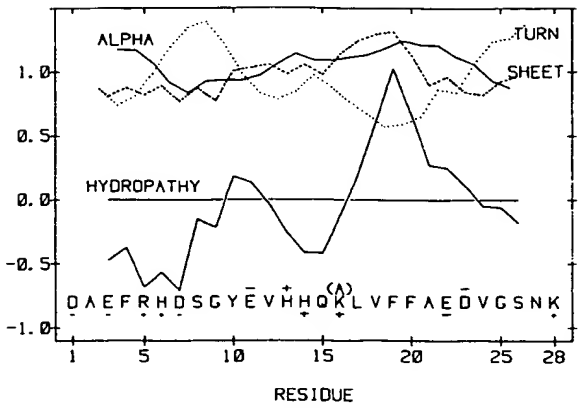

FIG. 1. Primary sequence and secondary structure analysis for β-(1–28). Predicted secondary structure for α-helix, β-pleated sheet, and β-turn (24), and hydropathic profile (26). Analysis using an alternative predictor (25) gave similar results. The primary amino acid sequence is shown just above the *x*-axis. The single substitution at residue 16 to produce [Ala[16]]β-(1–28) is indicated. Charged groups are denoted by the (+) and (−) signs, which are placed above or below the amino acid in the putative β-strand (residues 11–24) to indicate sidedness of the residue.

diffraction. The β-(18–28) peptide solution showed no precipitates even at 30 mg/ml, so it was lyophilized and examined by x-ray diffraction after vapor rehydration.

Electron Microscopy. For thin sectioning, samples were fixed in 2.5% (vol/vol) buffered glutaraldehyde (pH 7.1), post-fixed in 2% (wt/vol) OsO_4, dehydrated through a graded series of ethanol, *en bloc* stained with 5% (wt/vol) uranyl acetate, and embedded in Epon. Thin sections were counterstained with 3% (wt/vol) uranyl acetate and Reynold's lead citrate. For negative staining, aliquots were placed on carbon-coated, Formvar-covered copper grids, and stained with 2% (wt/vol) uranyl acetate. Samples were observed in a JEOL 100-S electron microscope at 80 kV. Calibration of the micrographs was carried out using tropomyosin paracrystals (kindly provided by C. Cohen, Brandeis University, Waltham, MA).

X-Ray Diffraction. Experiments were carried out using Ni-filtered and double-mirror focused CuK_α radiation from an Elliott GX-20 rotating anode generator (GEC Avionics, Hertfordshire, England) operated at 35 kV and 55 mA. Patterns were recorded on Kodak direct-exposure x-ray film. Specimen-to-film distances were 71.3 or 171.6 mm. Measurement of the diffraction spacings, integral linewidths, and arc lengths was carried out as described (23).

Antibodies. A rabbit antiserum was raised against a low molecular weight protein (4–7 kDa) that had been extracted with formic acid from AD amyloid filaments, purified by size-exclusion HPLC as described, and had the amino acid composition of the β protein (9). An antiserum was also produced to the synthetic β-(1–28) peptide (9). In sections of AD and normal, aged brain, both of these antisera label only plaque and vascular amyloid (9, 13). Inhibition assays were performed by incubating various concentrations of peptides with antiserum for 2 hr at room temperature and then testing by ELISA assay as described (22).

RESULTS AND INTERPRETATION

Predicted Conformation Is a β-Structure. From the primary sequence of the peptides we estimated the conformation indices for α-helix, β-helix, β-turn (24, 25), and hydrophobicity (26) (Fig. 1 *Upper*). High β-turn potentials are centered at residues 8 and 26. The intervening sequence shows a high potential for either β-conformation or α-helix. The proximal half of the region between the β-turns has three positive charges (His-13, His-14, and Lys-16), or two in [Ala[16]]β-(1–28), whereas the distal half has two negative charges (Glu-22 and Asp-23). The N-terminal region proximal to the first β-turn is hydrophilic and shows some α-helical potential.

Electron Microscopy Reveals Fibrils. Thin sections of β-(1–28) revealed assemblies that were fibrillar and aggregated into dense bundles (Fig. 2*a*) reminiscent of those comprising *in situ* and isolated amyloid cores (9). Fibril diameters were 60–80 Å, and lengths were 1200–1600 Å (Fig. 2 *a* and *b*); cross sections were occasionally visible and showed circular or oval profiles with electron-lucent cores. Negative staining of β-(1–28) (Fig. 2*c*) revealed periodic staining every 160 Å along some of the

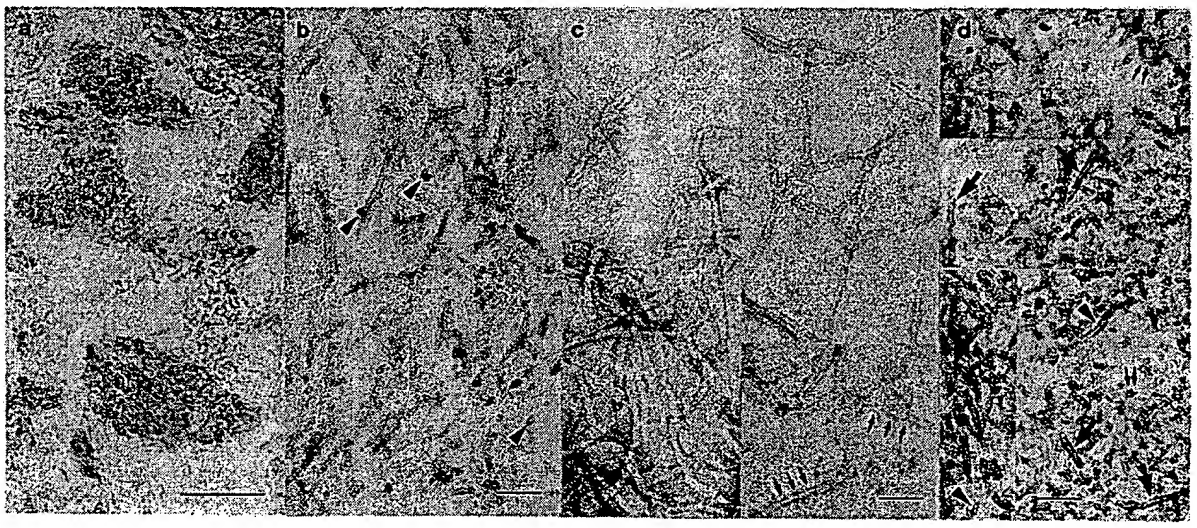

FIG. 2. Electron micrographs from synthetic β-protein homologues. (*a*) β-(1–28) in thin section at low magnification shows bundles of fibrils. (Bar = 1 μm.) (*b*) β-(1–28) in thin section at high magnification shows the individual fibrils. Oval, electron-lucent cross-sections of the fibrils are occasionally visible (arrowheads). (*c*) β-(1–28) fibrils, negatively stained, are periodically stained along some of their edges (arrows). Ends of fibrils appear to splay or unfurl. (*d*) [Ala[16]]β-(1–28) in thin sections at high magnification. An assortment of assemblies are seen, including linear (arrows), curvilinear (arrowheads), and rosette (double arrows) arrangements. (For *b*, *c*, and *d*, bar = 500 Å.)

fibrils. Aggregates of [Ala^{16}]β-(1–28) appeared amorphous at low magnification and, at high magnification, showed dense and disordered polymorphic assemblies (Fig. 2*d*). The linear assemblies had diameters similar to the β-(1–28) fibrils but did not associate in bundles.

Hydrated β-(1–28) and [Ala^{16}]β-(1–28) Give β-Type X-Ray Patterns. Diffraction patterns showed a sharp ring at 4.76-Å spacing and diffuse scatter in the 10-Å region (Fig. 3, curves a and b), features which are similar to those recorded from pellets of fibrils isolated from systemic amyloidoses (27, 28) and from enriched preparations of paired helical filaments and amyloid plaque cores from AD brain (23). These x-ray spacings indicate a β-pleated sheet conformation in which 4.76 Å is the distance between neighboring hydrogen-bonded polypeptide chains, and the scatter at ≈10 Å corresponds to the distance between pleated sheets (29).

Fiber Diffraction Pattern from Partially Dried β-(1–28). When allowed to dry, the β-(1–28) fibrils became oriented as shown by both electron microscopy and x-ray diffraction (Fig. 4). The accentuation of the hydrogen bond and intersheet reflections at right angles to one another demonstrated the cross-β-conformation in which the polypeptide chains lie normal to the axis of the fibril (29). Measurements of the arc length and breadth of the 4.76-Å reflection indicated that the fibrils were oriented within 20° of their average direction and that the coherence length along a fibril was 160–200 Å. The breadth of the 10.6-Å band indicated that the coherence length in the intersheet direction was nearly 30 Å, which corresponds at most to three pleated sheets.

The most prominent feature of the pattern was a very intense equatorial reflection at 71-Å spacing (Fig. 4 *Inset*). Higher orders of this spacing were observed as much weaker equatorial maxima (Fig. 4 and Table 1). The scatter centered at 71 Å corresponds to the first maximum of the square of a zero-order Bessel function calculated for a tubular cylinder having a diameter of 86 Å; it also could come from slab-like walls centered 71 Å apart. Since these calculated diameters are for unfixed material, we also recorded diffraction from a β-(1–28) sample after it was processed for electron microscopy (i.e., after fixing, dehydrating, and embedding). The embedded sample gave a strong equatorial reflection at 62-Å spacing that corresponds to a 75-Å diameter tubular cylinder or to slab-like walls separated by 62 Å. Both of these diameters are within the range of fibril diameters measured from electron micrographs.

To account for the subsequent intensity maxima on the equator, we computed transforms of step-function models. We

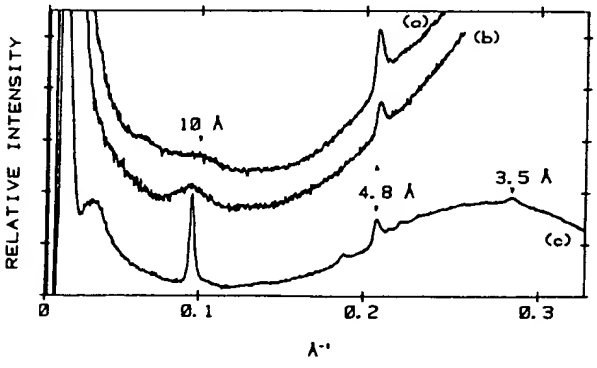

FIG. 3. Densitometer scans of x-ray patterns from hydrated gel of β-(1–28) (curve a), hydrated precipitate of [Ala^{16}]β-(1–28) (curve b), and vapor-hydrated lyophilized β-(18–28) (curve c). The positions of the interchain spacing (4.76 Å) and the intersheet spacing (≈10 Å) in curves a–c and of the chain repeat (3.5 Å) in curve c are indicated. The scans have been offset vertically for clarity.

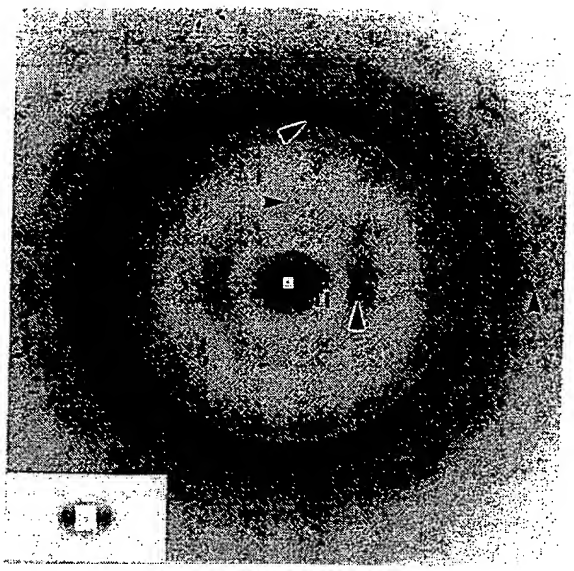

FIG. 4. X-ray diffraction pattern from β-(1–28) gel air-dried in a siliconized capillary. The beam was directed normal to the fiber axis, which is vertical. The strongly accentuated reflections along the meridian (4.76 Å; white-edged arrowhead) and equator (10 Å; white-edged arrowhead) indicate the cross-β-polypeptide conformation. The positions of selected other intensity maxima on the meridian (9.4 Å, arrowhead; 3.8 Å, arrow) and equator (24 Å and 18 Å, vertical bars; 3.3 Å, arrowhead) are indicated (see also Table 1). The small-angle region of the pattern, which contains the intense 71-Å spacing and which is overexposed in the top film, is shown in the *Inset* at a two-times enlargement of an underfilm.

took the peptide backbone thickness to be 4 Å and the intersheet space to be 6 Å (table 10.1 in ref. 29), and, ignoring the contribution of residual bound water, we calculated the electron-density levels of the backbone and of the side chains to be 0.441 electron per Å^3 and 0.289 electron per Å^3, respectively. Reasonable agreement was obtained between the observed and calculated positions of the intensity maxima (Table 1) for either the tubular cylinder model or the slab-walls model and in which the fibril walls consisted of pairs of β-pleated sheets.

Similarity to Diffraction Pattern of β-Keratin. The unit cell comprising the β-(1–28) fibril can be represented by a two-dimensional lattice having dimensions of 9.4 Å (hydrogen bond or fiber axis) and 6.6 Å (chain axis). Most of the observed spacings and intensities relating to this unit cell are similar to those reported for β-keratin (ref. 30; Table 1) and for certain heated and stretched globular proteins (34). In β-(1–28), the weak, equatorial arc at 3.3-Å spacing was identified as the unit translation along the polypeptide chain. This is close to the values suggested for parallel or antiparallel chains (35, 36). That the β-strands are antiparallel in β-(1–28) is shown by the weak but sharp reflection at 9.4 Å on the meridian.

Effect of Substituting Alanine for Lysine-16. Conditions that produced oriented β-(1–28) fibrils did not produce any orientation of the [Ala^{16}]β-(1–28) assemblies. The coherence length in the hydrogen-bonding direction was similar to that in β-(1–28), but in the intersheet direction it was about twice that of β-(1–28). Apparently the single substitution of the uncharged alanine for the positively charged lysine in the β-strand enhances the packing of the sheets.

Distal Half of β-Strand [β-(18–28)] Forms Extensive Intersheet Stacking. The major x-ray reflections from hydrated samples of lyophilized β-(18–28) were at 10.5 Å, 4.78 Å, and 3.53 Å spacings (Fig. 3, curve c), with the intersheet spacing not only the most prominent but also very sharp. The coherence length measured from the breadth of this reflection

Table 1. Principal x-ray spacings from synthetic peptide β-(1–28)

Comparison A: With step models			Comparison B: With β-keratin			
	d_c, Å					
d_o, Å	Slab walls	Tubular cylinder	d_o, Å	Index	d_c, Å	$d_{\beta k}$, Å
71.0 (*vsE*)	71.4	70.9	9.45 (*fM*)	(10)	9.52	
36.1 (*E*)	35.7	39.3	4.76 (*vsM*)	(20)	4.76	4.68 (*vs*)
24.5 (*mE*)	24.4	27.2	4.5* (*fM*)			
21 (*fM*)		20.9	4.3 (*fE*)			
17.6 (*fE*)	18.5		3.75† (*mM*)	(21)	3.86	3.80 (*s*)
12.3 (*sE*) ‡	13.9	13.6	3.3 (*fE*)	(02)	3.30	3.33 (*s*)
	11.8	11.8	2.76 (*vfE'*)	(22)	2.71	2.68 (*m*)
10.6 (*sE*) ‡	10.2	10.5	2.39 (*vfM*)	(40)	2.38	2.37 (*m*)
9.3 (*sE'*) ‡	8.9	9.3	2.24 (*vfM'*) §	(41)	2.24	2.23 (*m*)
			2.23 (*vfE*) §	(03)	2.20	—
			—	(13)	2.14	2.15 (*m*)
			2.05 (*vfE'*) §	(23)	2.00	2.00 (*s*)

d_o and d_c, observed and calculated Bragg spacing for β-(1–28); $d_{\beta k}$, observed spacing and relative intensity for β-keratin (30). In comparison A, models consisted of a pair of cross-β-pleated sheets either centered 71 Å from another pair ("slab-walls") or constituting the wall of an 86-Å-diameter tubular cylinder. In comparison B, indexing of the reflections for β-(1–28) were based on an orthogonal pseudo-unit cell of dimensions a = 9.52 Å (hydrogen-bonding direction) and b = 6.6 Å (chain direction), and for β-keratin the unit cell is orthogonal with a = 9.46 Å and b = 6.68 Å (30). Note that in β-keratin, the polypeptide chain axis is parallel to the fiber axis. *E*, equator; *M*, meridian; *E'*, off-equatorial; *M'*, off-meridional; (*vf, f, m, s, vs* indicate increasing intensity from very faint, faint, moderate, strong, and very strong.

*When indexed as (201) this reflection is consistent with a quarter-stagger arrangement of the sheets (31) with the *c*-axis along the intersheet direction and c = 21.2 Å.

†This spacing, which is probably off-meridional, has also been recorded in x-ray patterns from isolated systemic amyloids (32, 33).

‡Region of very broad scatter.

§This spacing and those at wider angles were measured from films recorded using a 26-mm camera.

was 250 Å, indicating stacks of ≈24 sheets. The broad band centered at ≈31-Å spacing may arise from the end-to-end packing of the 11-residue β-strands.

Antibody Recognition. Studies utilizing chemical modification of synthetic β-(1–28) indicate that Lys-16 is involved in the recognition of AD amyloid by antisera to either native or synthetic β protein (22). Using an ELISA assay we tested whether the three synthetic peptides could also inhibit the immunoreactivity of antisera to the HPLC-isolated, native amyloid protein. We found that the binding of antibody to amyloid was inhibited 80% by β-(1–28) and 60% by [Ala^{16}]β-(1–28), but not at all by β-(18–28). This suggests that the antigenic site in AD amyloid is exposed in both β-(1–28) and [Ala^{16}]β-(1–28) but absent in β-(18–28).

DISCUSSION

We have observed three different types of aggregates formed from three synthetic peptides that have sequence homologies with the amyloid β protein of AD. The synthetic peptide having the same sequence as the N-terminal 28 residues formed fibrils that share both antigenic and structural features with native AD amyloid filaments. These fibrils assembled spontaneously in aqueous solution, tended to pack loosely and in parallel bundles, became oriented during drying, and showed a cross-β-conformation. The synthetic peptide having a single change in the sequence, i.e., Lys-16 to Ala, formed a variety of assemblies including linear arrays and curvilinear clusters. These assemblies precipitated from solution, did not spontaneously orient, and showed a β-conformation. The synthetic peptide consisting of residues 18–28 of the β protein formed assemblies that showed extensive stacks of β-pleated sheets. The *in vitro* formation of amyloid-like fibrils from synthetic peptides with residues 12–28 and 1–28 with a Gln-11 substitution has also been reported (37).

As demonstrated by the β-protein homologues, the formation of amyloid-like fibrils clearly depends on sequence specificity and extent. What are the important determinants of these β-conformation structures? What promotes fibril formation, and what hinders it? The structures examined all showed about the same extent of hydrogen bonding between chains. Thus, while hydrogen bonding is required, it is not a specific determinant of fibrilogenesis. What differed most was the extent of intersheet packing: 2–3 sheets in β-(1–28),

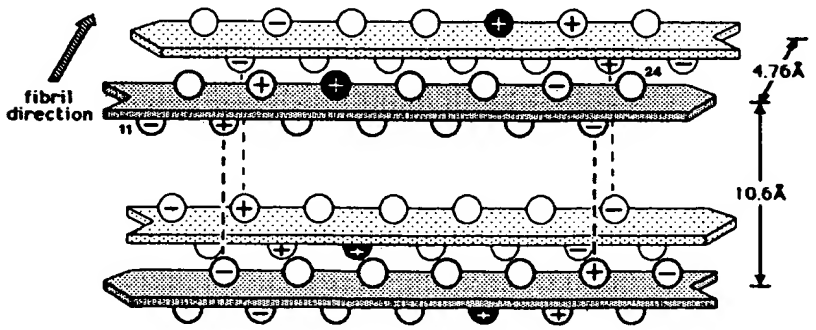

FIG. 5. Schematic model of a tetramer of the putative β-strand region (residues 11–24) of β-(1–28). The β-strands are shown antiparallel, as indicated by the x-ray spacing at 9.4 Å. The residues of the β-strand are represented by circles with the charges marked. The intersheet interaction shown is with the Lys-16 (dark circles) facing outward, as suggested by the structural and inhibition assay findings with [Ala^{16}]β-(1–28). Dashed lines indicate possible ion pairing.

5–6 sheets in [Ala[16]]β-(1–28), and ≈24 sheets in β-(18–28). It may be that the fibrils form only if the sequence of side chains promotes intersheet packing along one side and impedes it along the other side, perhaps due to steric hindrance or charge repulsion.

The side chains of apposed sheets may interact in either parallel or antiparallel fashion. In β-(1–28), the parallel mode is unlikely, due to repulsion between positively charged groups in the proximal region of the β-strand and between negatively charged groups in the distal region (see Fig. 1). In the antiparallel mode ion pairing would complement hydrophobic interactions and provide additional stability. Since substitution of lysine-16 by alanine promotes intersheet stacking, it may be that this side is exposed and not involved in intersheet interactions in β-(1–28). The exposed nature of this side is also supported by the finding that modification of lysine-16, either by derivatization (22) or by the substitution of alanine (this study), reduces its ability to inhibit antibody binding to amyloid. Fig. 5 shows a schematic tetramer of the putative β-strand portion of β-(1–28) that illustrates the proposed intersheet arrangement.

[Ala[16]]β-(1–28) did not form amyloid-like fibrils and bundles. The alanine substitution apparently allows other possible assemblies to form, due perhaps to the additional intersheet stacking at this side of the β-strand. In β-(18–28), the extensive stacking of sheets may be due both to ion pairing involving the terminal lysine and the glutamic and aspartic acid side chains and to hydrophobic interactions among the other side chains of the apposed, antiparallel strands. This would result in a staggered arrangement of the chains. The high water solubility of β-(18–28) compared to β-(1–28) and [Ala[16]]β-(1–28) suggests that the stabilizing interactions in β-(18–28) are mostly ionic.

We have demonstrated that a synthetic peptide identical in sequence to the first 28 amino acids of the 4-kDa amyloid protein can assemble *in vitro* into fibrils and bundles having very similar structural and antigenic features as those observed *in situ*. Our findings provide a basis for further *in vitro* modeling of amyloid fibril assembly, structure, and possibly also disruption of such assemblies. For example, the model system could be used to examine the effects of other amino acid substitutions in the amyloid protein, or post-translational modifications and various chain lengths of the putative precursor protein on amyloid fibril formation.

We thank Paul Tempst and Gary Gryam for helpful discussions on peptide synthesis, William P. McIntosh for photographic services, and Dr. Lee Makowski for critical comments on the manuscript. The research was supported in part by an Alzheimer Disease and Related Disorders Association Pilot Grant (D.A.K.); Grants NS-20824 (D.A.K.), NS-23023 (L.K.D.), AG-01307 (D.J.S.), and NS-20110 (D.J.S.) from the National Institutes of Health and an American Federation for Aging Research grant (L.K.D.). D.J.S. is a recipient of a Metropolitan Life Foundation Award for Medical Research. The work was carried out in facilities related to the Mental Retardation Research Center of Children's Hospital, and was supported by Core Grant HD-06276 from the National Institutes of Health.

1. Tomlinson, B. E. & Corsellis, J. A. N. (1984) in *Greenfield's Neuropathology*, eds. Adams, J. H., Corsellis, J. A. N. & Duchen, L. W. (Arnold, London), pp. 951–1025.
2. Narang, H. K. (1980) *J. Neuropathol. Exp. Neurol.* **39**, 621–631.
3. Merz, P. A., Wisniewski, H. M., Somerville, R. A., Bobin, S. A., Masters, C. L. & Iqbal, K. (1983) *Acta Neuropathol.* **60**, 113–124.
4. Glenner, G. G., Henry, J. H. & Fujihara, S. (1981) *Ann. Pathol.* **1**, 120–129.
5. Vanley, C. T., Aguilar, M. H., Kleinhenz, R. J. & Lagios, M. D. (1981) *Hum. Pathol.* **12**, 609–616.
6. Glenner, G. G. (1983) in *Biological Aspects of Alzheimer's Disease: Banbury Report*, ed. Katzman, R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 15, pp. 127–144.
7. Wong, C. W., Quaranta, V. & Glenner, G. G. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8729–8732.
8. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L. & Beyreuther, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4245–4249.
9. Selkoe, D. J., Abraham, C. R., Podlisny, M. B. & Duffy, L. K. (1986) *J. Neurochem.* **46**, 1820–1834.
10. Joachim, C. L., Morris, J. & Selkoe, D. J. (1986) *Neurology* **36**, Suppl. 1, 226.
11. Epstein, C. J. (1983) in *Biological Aspects of Alzheimer's Disease: Banbury Report*, ed. Katzman, R. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 15, pp. 169–182.
12. Tomlinson, B. E., Blessed, G. & Roth, M. (1968) *J. Neurol. Sci.* **7**, 331–356.
13. Selkoe, D. J., Bell, D., Podlisny, M. B., Cork, L. C. & Price, D. L. (1987) *Science* **235**, 873–877.
14. Glenner, G. G. & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* **120**, 885–890.
15. Glenner, G. G. & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* **122**, 1131–1135.
16. Masters, C. L., Multhaup, G., Simms, G., Pottgiesser, J., Martins, R. N. & Beyreuther, K. (1985) *EMBO J.* **4**, 2757–2763.
17. Bobin, S. A., Currie, J. R., Chen, M. C., Iqbal, K., Miller, D. L., Styles, J. J. & Wisniewski, H. M. (1987) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **46**, 2136 (abstr.).
18. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U. & Gajdusek, D. C. (1987) *Science* **235**, 877–880.
19. Robakis, N. K., Wisniewski, H. M., Jenkins, E. C., Devine-Gage, E. A., Houck, G. E., Yao, X.-L., Ramakrishna, N., Wolfe, G., Silverman, W. P. & Brown, W. T. (1987) *Lancet* **i**, 384–385.
20. Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A. P., St. George-Hyslop, P., VanKeuren, M. L., Patterson, D., Pagan, S., Kurnit, D. M. & Neve, R. L. (1987) *Science* **235**, 880–884.
21. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K.-H., Multhaup, G., Beyreuther, K. & Muller-Hill, B. (1987) *Nature (London)* **325**, 733–736.
22. Duffy, L. K., Kates, J., Podlisny, M. B., Walsh, R., Galibert, L. & Selkoe, D. J. (1986) *Protides Biol. Fluids Proc. Colloq.* **34**, 193–196.
23. Kirschner, D. A., Abraham, C. & Selkoe, D. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 503–507.
24. Chou, P. Y. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276.
25. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
26. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. (1984) *J. Mol. Biol.* **179**, 125–142.
27. Eanes, E. D. & Glenner, G. G. (1968) *J. Histochem. Cytochem.* **16**, 673–677.
28. Bonar, L., Cohen, A. S. & Skinner, M. M. (1969) *Proc. Soc. Exp. Biol. Med.* **131**, 1373–1375.
29. Fraser, R. D. B. & MacRae, T. P. (1973) *Conformation in Fibrous Proteins and Related Synthetic Polypeptides* (Academic, New York), pp. 179–246.
30. Fraser, R. D. B. & MacRae, T. P. (1962) *J. Mol. Biol.* **5**, 457–466.
31. Geddes, A. J., Parker, K. D., Atkins, E. D. T. & Beighton, E. (1968) *J. Mol. Biol.* **32**, 343–358.
32. Termine, J. D., Eanes, E. D., Ein, D. & Glenner, G. G. (1972) *Biopolymers* **11**, 1103–1113.
33. van Andel, A. C. J., Hol, P. R., van der Maas, J. H., Lutz, E. T. G., Krabbendam, H. & Gruys, E. (1986) in *Amyloidosis*, eds. Glenner, G. G., Osserman, E. F., Benditt, E. P., Calkins, E., Cohen, A. S. & Zucker-Franklin, D. (Plenum, New York), pp. 39–48.
34. Senti, F. R., Eddy, C. R. & Nutting, G. C. (1943) *J. Am. Chem. Soc.* **65**, 2473.
35. Pauling, L. & Corey, R. B. (1953) *Proc. Natl. Acad. Sci. USA* **39**, 251–271.
36. Fraser, R. D. B., MacRae, T. P., Parry, D. A. D. & Suzuki, E. (1969) *Polymer* **10**, 810–826.
37. Castano, E. M., Ghiso, J., Prelli, F., Gorevic, P. D., Migheli, A. & Frangione, B. (1986) *Biochem. Biophys. Res. Commun.* **141**, 782–789.

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 5355–5358, June 1994
Biochemistry

Glutamine repeats as polar zippers: Their possible role in inherited neurodegenerative diseases

MAX F. PERUTZ, TONY JOHNSON, MASASHI SUZUKI, AND JOHN T. FINCH

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Contributed by Max F. Perutz, February 25, 1994

ABSTRACT **Four inherited neurodegenerative diseases are linked to abnormally expanded repeats of glutamine residues in the affected proteins. Molecular modeling followed by optical, electron, and x-ray diffraction studies of a synthetic poly(L-glutamine) shows that it forms β-sheets strongly held together by hydrogen bonds. Glutamine repeats may function as polar zippers, for example, by joining specific transcription factors bound to separate DNA segments. Their extension may cause disease either by increased, nonspecific affinity between such factors or by gradual precipitation of the affected proteins in neurons.**

Four inherited neurodegenerative diseases are linked to abnormally expanded repeats of glutamines near the N termini of the affected proteins (for reviews, see refs. 1 and 2). They are Huntington disease (HD); spinal and bulbar muscular atrophy (SBMA), also known as Kennedy disease; spinocerebellar ataxia type 1 (SCA1); and dentatorubral-pallidoluysian atrophy (DRPLA) (3–8). All four diseases become more severe and begin earlier the longer the glutamine repeats. The repeats tend to lengthen in successive generations of affected individuals, especially in male transmission. So far no molecular function has been proposed for the glutamine repeats. We suggest that they act as polar zippers, joining protein molecules together similarly to the way leucine zippers join the transcription factors c-Jun and c-Fos.

The gene for the HD protein contains an open reading frame for >3100 amino acid residues. Its amino acid sequence shows no homology with any known protein. The glutamine-rich segment starts at residue 18. In healthy individuals its length varies between 16 and 33 residues; in those afflicted by HD it may vary between 35 and >100 residues. It is followed by an almost continuous stretch of 29 prolines. Poly(L-proline) forms a rigid helix with threefold screw symmetry and an axial repeat of 3.1 Å per residue (9). Its position near the N terminus of the HD protein implies that the glutamine repeat is mounted at the end of either a 90-Å-long stalk or, if an intervening peptide with the sequence QAQPLLPQPQ divides it into two, at the end of two consecutive, >30-Å-long stalks: a bizarre structure of a kind not encountered before.

SBMA is linked to the expansion of a glutamine-rich segment in the androgen receptor. This is a transcription factor made of a 920-residue chain with one DNA-binding C-terminal domain. Starting at residue 58, healthy individuals' receptors have a sequence of 15–31 glutamines. In SBMA patients this is expanded to 40–62 (4, 5). An androgen receptor with an expanded glutamine repeat trans-activated an androgen-responsive reporter gene more weakly than a normal receptor (10). No complete amino acid sequence of the protein responsible for SCA1 has been published. The length of its glutamine repeat varies from 25 to 36 residues in normal individuals from 43 to 81 residues in SCA1 patients (6).

Structure of Poly(L-glutamine). What could be the structure of poly(L-glutamine)? Perutz *et al.* (11) found several proteins containing repetitive sequences of polar residues. Molecular modeling showed these to be capable of linking β-strands together into sheets or barrels by networks of hydrogen bonds between their main-chain amides and between their polar side chains. Perutz *et al.* called these sequences polar zippers. One such sequence was a continuous stretch of up to 65 glutamines in the female sterile homeotic protein of *Drosophila* (12). Fig. 1 shows a computer-generated, stereochemically satisfactory model of two pairs of antiparallel β-strands of poly(L-glutamine) linked together by hydrogen bonds between their main-chain and side-chain amides.

When the importance of glutamine repeats in human disease became apparent, we decided to test the validity of this model experimentally, but this was difficult because poly(L-glutamine) is insoluble in water. To make it soluble, we synthesized a peptide with the sequence Asp_2-Gln_{15}-Lys_2.* Filtration of the peptide through Beckman Ultraspherogel SEC 2000 showed two fractions, one corresponding to the molecular weight of a monomer and the other to a broad distribution of aggregates of molecular weights in the range of hundreds of thousands. We investigated the conformation of the peptide by ultraviolet circular dichroism (CD). To avoid interactions between COO^- and NH_3^+ groups, we measured the CD spectra in solutions at pH 3.0 and 2.0, where the carboxylates are protonated. Peptides forming α-helices, β-sheets, or random coils give the different CD spectra shown in Fig. 2*a*. Fig. 2*b* shows the CD spectra of our peptide in three different solvents to be of the β-sheet type, even though one of the solvents, trifluoroethanol, normally induces formation of α-helices. Molar residue ellipticity at 197 nm was independent of peptide concentration, suggesting that the monomers form hydrogen-bonded hairpins and that the aggregates are made of tightly linked β-sheets. The spectra were qualitatively the same at pH 7.0 and 3.0. At pH 7.0 the peptide gradually precipitated. When viewed under

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HD, Huntington disease; SBMA, spinal and bulbar muscular atrophy; SCA1, spinocerebellar ataxia type 1; DRPLA, dentatorubral-pallidoluysian atrophy.

*The peptide was prepared by continuous-flow solid-phase synthesis by fluorenylmethoxycarbonyl chemistry on an automated synthesizer (NovaSyn Crystal) employing standard protocols (13, 14). On completion of the synthesis, the free N terminus was acetylated with acetic anhydride (20 equivalents) for 30 min. The peptide was cleaved from the Rink amide linker with trifluoroacetic acid/phenol/triethylsilane (23 ml/1 g/1 ml for 500 mg of peptide-resin assembly) and purified to homogeneity by preparative HPLC (Vydac C_8 column) (13, 14). Analysis of the purified material showed the following. Amino acid analysis: expected, $Asp_2Glx_{15}Lys_2$; found, $Asp_{1.91}Glx_{15.86}Lys_{2.00}$. Analytical reversed-phase HPLC (Aquapore RP-300, C_8 column): retention time, 12.62 min (>98%). Capillary zone electrophoresis (Beckman P/ACE system 2050; 75 mm × 50 cm capillary; 100 mM phosphate buffer, pH 2.5; 30 kV, 30°C); retention time, 15.19 min (>98%). Electrospray mass spectrometry: expected for Ac-$(Asp)_2$-$(Gln)_{15}$-$(Lys)_2$-NH_2, 2467.54; found, 2467.25, 2490.85 $(M+Na)^+$, 2505.63 $(M+K)^+$.

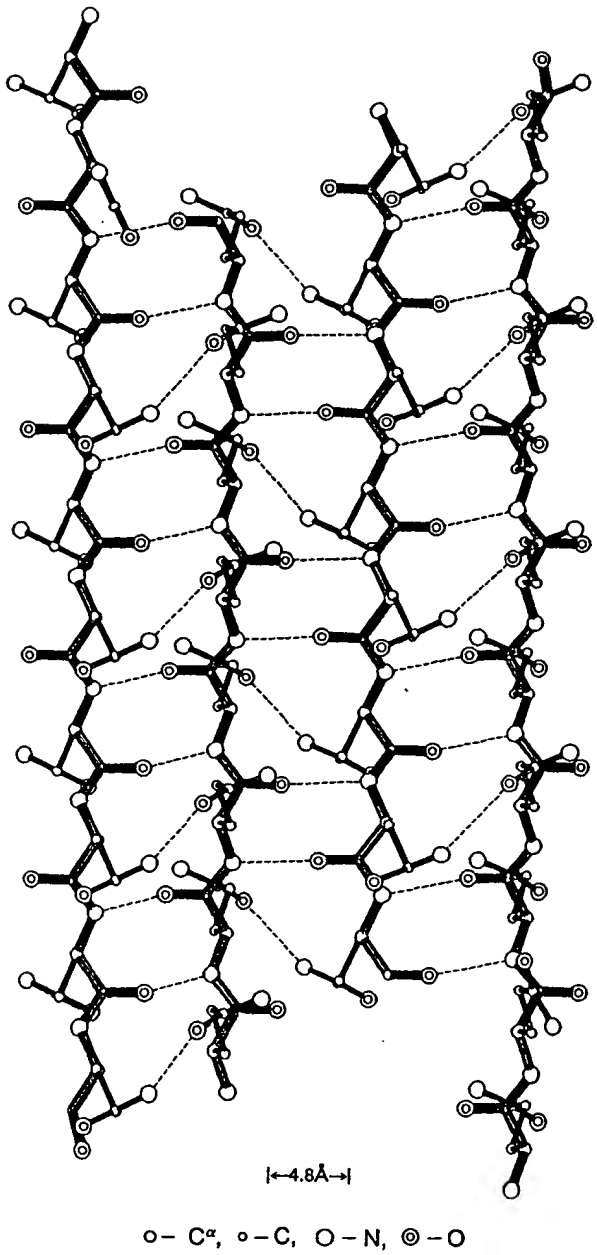

FIG. 1. Computer-generated structure of two paired antiparallel β-strands of poly(L-glutamine) linked together by hydrogen bonds between the main-chain and side-chain amides.

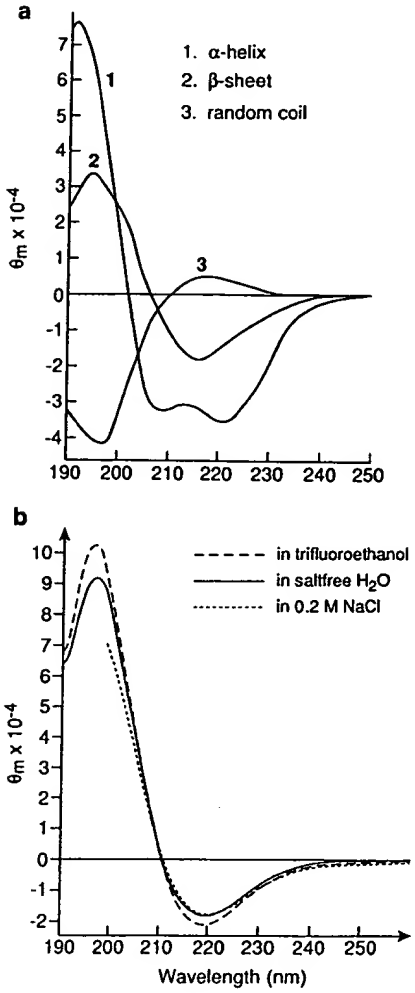

FIG. 2. (*a*) Standard CD spectra of α-helix, β-sheet, and random-coil structures of poly(L-lysine). (Reproduced by permission from ref. 15; copyright 1969, American Chemical Society.) (*b*) CD spectra of Asp_2-Gln_{15}-Lys_2 at 0.3 mg/ml in various media. The spectra were measured with a dichrograph (Jobin-Yvon CD6) at 20°C. Water and 0.2 M NaCl were adjusted to pH 3.0 by addition of trifluoroacetic acid. The pH of 95% 2,2,2-trifluoroethanol was 2.0.

the polarizing microscope, a suspension of the precipitate looked clear at first, but as the solvent between slide and coverslip began to evaporate, birefringence developed at the boundaries between air and water. The birefringence was positive, with the slow direction parallel to the meniscus. This suggested the presence of submicroscopic rod-shaped particles which surface tension had aligned parallel to the meniscus. We therefore examined the precipitate under the electron microscope. Fig. 3*a* shows that it consists of worm-like 70- to 120-Å-thick particles of varying length. An electron diffraction picture of a small clump of the dried particles (not shown) was dominated by a sharp ring of 4.8-Å spacing, diagnostic of the distance between neighboring polypeptide chains in β-sheets (16, 17). In addition, the picture contained faint rings at spacings of 8.4 and 4.2 Å. We next took an x-ray diffraction picture of the wet particles spun into a quartz capillary (not shown). This contained the same sharp rings as the electron diffraction picture plus another faint one at 2.8-Å spacing. A suspension of the particles was then dried on a glass slide. The dried film was lifted off and an x-ray diffraction picture was taken of it with the x-ray beam parallel to the plane of the film (Fig. 3). The picture exhibits a fiber diagram of the cross-β type with the 8.4-, 4.8-, and 4.2-Å reflections on the meridian and reflections at 3.6 and 3.2 Å on the equator. The 2.8-Å reflection lies beyond the rim of this picture. The equatorial reflection at 3.2 Å corresponds to the axial repeat per residue in a pleated β-sheet (16, 17). The reflection at 3.6 Å corresponds to the axial repeat per residue in a fully extended chain, which is not normally observed. Alternatively, it could represent a higher order of a long repeat that is not apparent in the rest of the picture. The strong meridional reflection at 4.8 Å and the equatorial one at 3.2 Å are indicative of a β-sheet with the chain direction

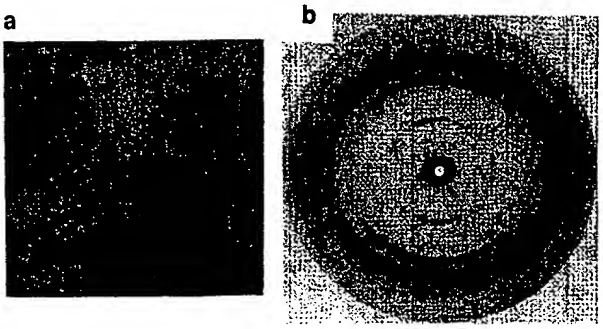

FIG. 3. (*a*) Electron micrograph of particles of Asp_2-Gln_{15}-Lys_2 formed in an aqueous solution at pH 7.0 and negatively stained with 1% uranyl acetate (× 90,000). (*b*) X-ray diffraction photograph (Cu K_α radiation; film distance, 37 mm) of the same particles oriented vertically, showing the dominant meridional reflection at 4.8 Å and the equatorial ones at 3.6 and 3.2 Å.

normal to the fiber axis. Molecular models show the spacing between successive β-sheets of poly(L-glutamine) hydrogen-bonded to each other by their amide side-chains to be about 17 Å, which suggests that the equatorial reflections at 8.4 and 4.2 Å and the diffraction ring at 2.8 Å could represent the second-, fourth-, and sixth-order reflections from parallel β-sheets stacked 16.8 Å apart. Odd orders may be weakened by the side-chain amides being stacked halfway between the main-chain amides. A likely structure suggested by the fiber diagram would be β-sheets rolled round each other. All these experimental results obtained from the 15-mer L-glutamine peptide are consistent with the polar zipper structure shown in Fig. 1.

Normal Functions of Glutamine Repeats. What possible function could the glutamine-rich segments in the HD, SBMA, SCA1, and SDRPLA proteins have? Long repeats of glutamines occur in many proteins, especially among transcription factors such as the homeobox proteins of *Drosophila*. Possible functions of glutamine repeats have been studied in the two homeoproteins coded for by the Abdominal B gene of *Drosophila*. Both proteins contain a C-terminal domain which includes the DNA-binding segment. One of them contains an additional N-terminal domain of 224 residues, 32% of which are glutamines. Deletion of part of this sequence did not diminish affinity for the homeoprotein-binding DNA segment, but it halved transcriptional activity. Conversely, splicing a distant part of the glutamine-rich domain to an otherwise inactive C-terminal domain restored transcriptional activity, which showed that the glutamine-rich domain did serve a necessary function, but without a hint as to its possible mechanism (18).

Does it serve a function in humans? Next to nothing is known as yet about the function of the glutamine repeats in the proteins affected by the three diseases, but some clues have emerged from studies of the human transcription factor Sp1. This is a trans-activator of gene expression which binds to G+C-rich regions in the promoters of several cellular and viral genes, including that of simian virus 40. The C-terminal fragment of Sp1 is a single chain of 696 residues with a glutamine-rich segment between residues 260 and 391 and with three zinc fingers beyond residue 540 (19). Courey and Tijan (20) designed an *in vivo* translation assay to study the function of the glutamine-rich segment. A C-terminal fragment of Sp1 containing the zinc fingers was transcriptionally inactive but became activated after a fragment containing residues 369–391 had been spliced onto it. This had the sequence PGNQVSWQTLQLQNLQVQNPQAQ; it contained 8 glutamines, 3 asparagines, a serine, and a threonine; these are all residues with polar side chains capable of forming hydrogen bonds with complementary polar side chains in neighboring β-strands (11). Alternatively, the inactive C-terminal fragment could be activated by splicing to it the glutamine-rich segment of the Antennapedia protein of *Drosophila* (21). The same authors next constructed a reporter plasmid with a single G-C box close to the initiation site of transcription, and six G-C boxes 1.7 kb downstream. Binding of Sp1 to these distantly placed G-C boxes stimulated the weak activation induced by the binding of Sp1 to the close G-C box. A similar stimulation was produced by the addition of a truncated Sp1 from which the zinc fingers had been deleted but which included the glutamine-rich segment (21). These experiments demonstrated that interaction between the glutamine-rich segments of Sp1 molecules bound to widely separated DNA segments enhanced transcription.

On the other hand, no function has so far been found for a stretch of 38 consecutive glutamines in the human TATA box-binding factor TFIID, but its conservation in the mouse factor argues in favor of its fulfilling an essential function (22, 23).

Molecular Pathology of Glutamine Repeats. Dominant transmission of HD and SCA1 argues in favor of these diseases being due to gains rather than losses of functions. SBMA is a chromosome X-linked recessive disease that manifests itself only in males. In females X-inactivation produces somatic cell mosaicism in which only the abnormal gene is likely to be expressed in a proportion of cells and only the normal gene in the remainder. The neural network is thought to contain enough redundancy so that inactivation of those neurons in which the abnormal gene is expressed does not matter. Loss of function is excluded as a cause of SBMA by the absence of SBMA symptoms in patients suffering from deletion of the androgen receptor gene. Mhatre *et al.* (10) have therefore suggested that the gain of function in each of the three diseases consists in an "aberrant transregulatory activity."

One possibility is that glutamine repeats joining specific complementary proteins as part of normal transcriptional regulation acquire excessively high affinities for each other or acquire nonspecific affinities for other regulatory proteins when they become too long. Alternatively, and perhaps more probably in the light of our results, extensions of their glutamine repeats may cause the affected proteins to agglomerate and precipitate in neurons; symptoms may set in when these precipitates have reached a critical size or have resulted in a critical number of neural blocks. This would explain better why symptoms appear earlier in life and become more severe the longer the extension of the glutamine repeats and why the main histological manifestation of HD consists in neural degeneration. On the other hand, immunostaining of Purkinje cells and cells in the human frontal cerebral cortex with an antibody against the HD protein showed no difference between cells from a normal individual and those from a HD patient (24).

1. Martin, J. B. (1993) *Science* **262**, 674–676.
2. Ross, A. C., McInnis, M. G., Margolis, R. L. & Li, S.-H. (1993) *Trends Neurosci.* **16**, 254–259.
3. The Huntington's Disease Collaborative Group (1993) *Cell* **72**, 971–983.
4. La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fischbeck, K. H. (1991) *Nature (London)* **352**, 77–79.
5. La Spada, A. R., Roling, D. B., Harding, E. A., Warner, C. L., Spiegel, R., Hausmanova-Petrusewicz, I., Yee, W.-C. & Fischbeck, K. H. (1992) *Nature Genet.* **2**, 301–304.
6. Orr, H. T., Chung, M.-y., Banfi, S., Kwiatkowski, Jr., T. J., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P. W. & Zoghbi, H. Y. (1993) *Nature Genet.* **4**, 221–226.
7. Matilla, T., Volpini, V., Genis, D., Rosell, J., Corral, J., Davalos, A., Molins, A. & Estivill, X. (1993) *Hum. Mol. Genet.* **2**, 2123–2128.
8. Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S.,

Endo, K., Takahashi, H., Kondo, R., Ishikawa, A., Hayashi, T., Saito, M., Tomoda, A., Miike, T., Naito, H., Ikuta, F. & Tsuji, S. (1994) *Nature Genet.* **6**, 9–13.
9. Cowan, P. M. & McGavin, S. (1955) *Nature (London)* **176**, 501–503.
10. Mhatre, A. N., Trifiro, M. A., Kaufman, M., Kazemi-Esfarani, P., Figlewicz, D., Rouleau, G. & Pinsky, L. (1993) *Nature Genet.* **5**, 184–188.
11. Perutz, M. F., Staden, R., Moens, L. & De Baere, I. (1993) *Curr. Biol.* **3**, 249–253.
12. Haynes, S. R., Mozer, B. A., Bhatin-Dey, N. & Dawid, I. B. (1989) *Dev. Biol.* **134**, 246–257.
13. Atherton, E. & Sheppard, R. C. (1988) *Solid Phase Peptide Synthesis: A Practical Approach* (Oxford Univ. Press, Oxford).
14. Dryland, A. & Sheppard, R. C. (1988) *Tetrahedron* **44**, 859.
15. Greenfield, N. & Fasman, G. D. (1969) *Biochemistry* **8**, 4108–4116.
16. Struther-Arnott, S., Dover, S. D. & Elliott, A. (1967) *J. Mol. Biol.* **30**, 202–208.
17. Ashida, T., Tanaka, I. & Yamane, T. (1981) *Int. J. Pept. Protein Res.* **17**, 322–329.
18. Ali, N. & Bienz, M. (1991) *Mech. Dev.* **35**, 55–64.
19. Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tijan, R. (1987) *Cell* **51**, 1079–1090.
20. Courey, A. J. & Tijan, R. (1988) *Cell* **55**, 887–898.
21. Courey, A. J., Holtzman, D. A., Jackson, S. P. & Tijan, R. (1989) *Cell* **59**, 827–836.
22. Kao, C. C., Lieberman, P. M., Schmidt, M. C., Zhou, Q., Pei, R. & Berk, A. J. (1990) *Science* **248**, 1646–1650.
23. Zhou, Q., Boyer, J. G. & Berk, A. J. (1993) *Genes Dev.* **7**, 180–187.
24. Hoogeveen, A. T., Willemsen, R., Meyer, N., de Rooij, K. E., Roos, R. A. C., van Ommen, G.-J. B. & Galjaard, H. (1993) *Hum. Mol. Genet.* **2**, 2069–2073.

We propose that the only way to interpret all these measurements is to assume that in the 'paramagnetic state' the observed lattice distortions and ferromagnetic clusters are caused by the same entity—the magnetic polaron—which is responsible for the giant magnetoresistance effects. From the measurements at zero field we have shown that the two ingredients of the proposed magnetic polaron, a lattice distortion and a ferromagnetic cluster, are simultaneously present at temperatures of T_c–$2T_c$. One could think that they are uncorrelated and exist independently. However, the measurements of the change in volume and SANS intensity under magnetic field indicate that the ferromagnetic cluster is associated with the lattice distortion (small polaron). From these experiments, we conclude that both effects are linked. This is deduced from the way the curves of the volume lattice distortion (Fig. 2a) and the SANS intensity (Fig. 2b) superimpose under fixed magnetic field and, more importantly, the way the volume and the SANS intensity behave at fixed temperature versus the applied magnetic field (Fig. 3a and b). The data in Fig. 3a and b are easily understood considering that the magnetic polarons retain their structure unchanged up to ~1.5 T (at $1.1T_c$); the lattice distortion and the associated ferromagnetic cluster remain unchanged (no variation in either the volume or the SANS intensity). For larger magnetic fields, a tendency towards carrier delocalization and an increase in the size of the associated ferromagnetic cluster occurs.

The observed field effect on the character and degree of localization of the charge carriers provides evidence for the close relation between the existence of these entities (referred to as 'magnetic polarons') and the giant magnetoresistive effect. We would like to put emphasis on this aspect considering the result shown in Fig. 3c. The field-dependence of the resistivity is very similar to that found for ω and the SANS intensity. As a consequence, the mechanism for the giant magnetoresistance effect is related to the existence and nature of the magnetic polarons. A simple description of the observed behaviour should consider the polaronic lattice effect provided by a narrow polaronic band ('small polaron', that is, strong electron–phonon interaction) along with the magnetic character of the carrier (ferromagnetic cluster) which would make possible a broadening of the polaronic band under the presence of an applied magnetic field. Further theoretical developments on the giant magnetoresistance effect in these manganese perovskites should consider these ideas. □

Received 22 October 1996; accepted 4 February 1997.

1. Hellemans, A. *Science* **263**, 880–881 (1996).
2. Asamitsu, A., Moritomo, Y., Tomioka, Y., Arima, T. & Tokura, Y. *Nature* **373**, 407–409 (1995).
3. Zhao, G., Conder, K., Keller, H. & Müller, K. A. *Nature* **381**, 676–678 (1996).
4. Zhou, J.-S., Archibald, W. & Goodenough, J. B. *Nature* **381**, 770–772 (1996).
5. Röder, H., Zhang, Jun & Bishop, A. *Phys. Rev. Lett.* **76**, 1356–1359 (1996).
6. Kusters, R. M., Singleton, J., Keen, D. A., McGreevy, R. & Hayes, W. *Physica B* **155**, 362–365 (1989).
7. Von Helmolt, R., Wecker, J., Holzapfel, B., Schultz, L. & Samwer, K. *Phys. Rev. Lett.* **71**, 2331–2333 (1993).
8. Ibarra, M. R., Algarabel, P. A., Marquina, C., Blasco, J. & García, J. *Phys. Rev. Lett.* **75**, 3541–3544 (1995).
9. Kim, K. H., Gu, J. Y., Choi, H. S., Park, G. W. & Noh, T. W. *Phys. Rev. Lett.* **77**, 1877–1880 (1996).
10. Lynn, J. W. *et al. Phys. Rev. Lett.* **76**, 4046–4049 (1996).
11. De Teresa, J. M. *et al. Phys. Rev. B* **54**, 1187–1193 (1996).
12. Sun, J. Z., Krusin-Elbaum, L., Gupta, A., Gang, Xiao & Parkin, S. S. P. *Appl. Phys. Lett.* **69**, 1002–1004 (1996).
13. Holstein, T. *Ann. Phys.* **8**, 343–389 (1959).
14. Kasuya, T. & Yanase, A. *Rev. Mod. Phys.* **40**, 684–696 (1968).
15. Von Molnar, S. *et al. Phys. Rev. Lett.* **51**, 706–709 (1983).
16. Dietl, T. in *Semimagnetic Semiconductors and Diluted Magnetic Semiconductors* (eds Averous, M. & Balkanski, M.) 83 (Plenum, New York, 1951).
17. Wolf, P. A. in *Semiconductors and Semimetals* Vol. 25 (eds Furdyna, J. K. & Kossut, J.) 413–454 (Academic, New York, 1988).
18. Zener, C. *Phys. Rev.* **82**, 403–405 (1951).
19. Glatter, O. & Kratky, O. (eds) *Small Angle X-ray Scattering* (Academic, London, 1983).
20. Tanaka, J. & Mitsuhashi, T. *J. Phys. Soc. Jpn* **53**, 24–32 (1984).
21. Radaelli, P. G., Marezio, M., Hwang, H. Y., Cheong, S.-W. & Battlog, B. *Phys. Rev. B* **54**, 8992–8995 (1996).

Acknowledgements. We thank J. Stankiewizc for discussions. This work was supported by the Spanish DGICYT.

Correspondence and requests for materials should be addressed to M.R.I. (e-mail: ibarra@posta.unizar.es).

Responsive gels formed by the spontaneous self-assembly of peptides into polymeric β-sheet tapes

A. Aggeli*, M. Bell*, N. Boden*†, J. N. Keen*‡, P. F. Knowles*‡, T. C. B. McLeish*§, M. Pitkeathly‖ & S. E. Radford*‡

* *Centre for Self-Organising Molecular Systems, The University of Leeds, Leeds LS2 9JT, UK*
† *School of Chemistry,* ‡ *Department of Biochemistry & Molecular Biology, and* § *Department of Physics, The University of Leeds, Leeds LS2 9JT, UK*
‖ *Oxford Centre for Molecular Sciences and New Chemistry Laboratory, The University of Oxford, South Parks Road, Oxford OX1 3QT, UK*

Molecular self-assembly is becoming an increasingly popular route to new supramolecular structures and molecular materials[1–7]. The inspiration for such structures is commonly derived from self-assembling systems in biology. Here we show that a biological motif, the peptide β-sheet, can be exploited in designed oligopeptides that self-assemble into polymeric tapes and with potentially useful mechanical properties. We describe the construction of oligopeptides, rationally designed or based on segments of native proteins, that aggregate in suitable solvents into long, semi-flexible β-sheet tapes. These become entangled even at low volume fractions to form gels whose viscoelastic properties can be controlled by chemical (pH) or physical (shear) influences. We suggest that it should be possible to engineer a wide range of properties in these gels by appropriate choice of the peptide primary structure.

Our initial studies focused on a 24-residue peptide K24, whose primary structure (NH_2-Lys-Leu-Glu-Ala-Leu-Tyr-Val-Leu-Gly-Phe-Phe-Gly-Phe-Phe-Thr-Leu-Gly-Ile-Met-Leu-Ser-Tyr-Ile-Arg-COOH) is related to the transmembrane domain of the IsK protein[8]. This peptide, and a longer version of it (K27; ref. 9), readily form β-sheet structures in lipid bilayers, suggesting they would be good candidates for formation of β-sheet tapes in amphiphilic solvents. Indeed, we find that solutions of K24 in solvents such as methanol or 2-chloroethanol produce transparent viscoelastic gels at concentrations above 0.002 volume fraction (~3 mg ml^{-1}), stable up to the boiling point of the solvent. Infrared (IR; Fig. 1a) and circular dichroism (CD; Fig. 1b) spectra show that in methanol the peptide molecules have a predominantly antiparallel β-sheet conformation. The wide-angle X-ray diffraction of the gel reveals the presence of a 4.7 Å structural periodicity, consistent with the side-by-side separation expected for β-strands in β-sheet structures[10]. Transmission electron microscopy (TEM; Fig. 1c) reveals the presence of distinct, tape-like structures, ~8 nm wide (the length of the K24 molecule in the β-strand conformation) and with apparent lengths typically in excess of 0.1 μm.

The rheological properties of the gels provide further insight into their structure. The response to small-strain oscillatory shear (Fig. 2a) is typical of highly entangled polymer gels[11,12]. From the elastic modulus G'^0_N we calculate, assuming entropic rubber elasticity[11,12], a mesh size of 10–100 nm for gels with peptide volume fractions 0.03–0.003 (see Methods for details). The stress response is seen to be linear up to strains of 230% (Fig. 2b). This permits an estimate of a lower limit of 13 nm or greater for the persistence length of the tape, indicative of a moderately rigid polymer, consistent with the intrinsic rigidity of β-sheet structures. We have deduced from the rheological data (see Methods) an upper limit of 0.7 nm for the thickness of a tape, consistent with tapes a single molecule

thick. Atomic force microscopy also reveals tapes with thickness of 0.5–1 nm.

To elucidate the molecular interactions governing the self-assembly of peptides into β-sheet tapes, we have explored the behaviour of K24 as a function of the polarity ϵ_r and hydrogen bonding ability α of the solvent[13]. Our observations are represented as a 'phase diagram' in Fig. 3a. In solvents such as 1,1,1,3,3,3-hexafluoroisopropanol (HFIP) having $\alpha \geq 1.5$, clear, low-viscosity solutions comprising mixtures of helices and random coils are obtained, as shown by IR spectra (Fig. 1a) and far-ultraviolet CD spectra (Fig. 1b). By contrast, transparent, thermostable, self-supporting gels are formed in moderately polar solvents, such as methanol, with $25 < \epsilon_r < 68$, $\alpha < 1.5$, and peptide concentrations above 0.002–0.004 volume fraction (3–6 mg ml^{-1}). In less polar solvents ($15 < \epsilon_r < 25$) and in solvents with $\epsilon_r > 68$, the β-sheet tapes are not sufficiently soluble to form gels. Below $\epsilon_r \approx 15$, the proportion of peptide in the helical form increases. Addition of the surfactant SDS to the peptide gel formed in methanol increases the stability of the α-helix with respect to the β-sheet structure to such an extent that a simple newtonian fluid is obtained, demonstrating that the β-sheet structure is in part stabilized by attractive forces between hydrophobic side-chains in polar solvents[14].

To impart gelation to more polar solvents, β-sheet tapes with hydrophilic surfaces are a prerequisite. This is demonstrated by the

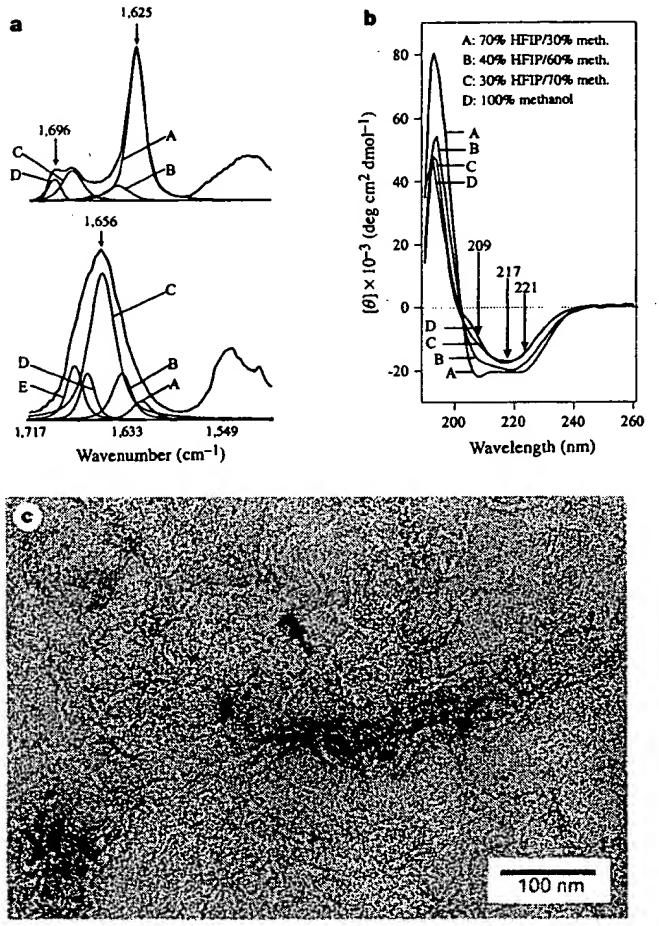

Figure 1 **a**, Band-fitted Fourier-transform IR spectra of amide I (1,700–1,600 cm^{-1}) and amide II (1,600–1,500 cm^{-1}) bands of K24 at volume fraction ~0.004 (~6 mg ml^{-1}). Component bands are labelled A–D. Top, gel in 100% methanol. Assignment of component bands: A 1,625 cm^{-1}, (maximum of amide I), β-sheet; B 1,641 cm^{-1}, random coil; C 1,680 cm^{-1}, residual trifluoroacetic acid (TFA); D 1,696 cm^{-1}, antiparallel β-sheet[33]. The half-height bandwidth of the major component at 1,625 is 17 cm^{-1}. The peak in the amide II region at 1,530 cm^{-1} is also consistent with a β-conformation. Bottom, spectrum of a clear, fluid solution of K24 in 90% 1,1,1,3,3,3-hexafluoroisopropanol (HFIP)/10% methanol. Assignment of component bands: A 1,620 cm^{-1} and B 1,638 cm^{-1}, β-sheet; C 1,655 cm^{-1} (maximum of amide I), α-helix; D 1,667 cm^{-1}, turns; E 1,678 cm^{-1}, TFA. The half-height bandwidth of the major component at 1,655 cm^{-1} is 26 cm^{-1}. The peak in the amide II band at ~1,545 cm^{-1} is also consistent with the predominance of helical conformation. The amide II bands are wider than the amide I bands, as expected[22]. Amide II bands are slightly distorted due to imperfect subtraction of solvent spectrum. **b**, Far-UV circular dichroism spectra of K24 in HFIP/methanol mixtures at peptide volume fraction ~0.00006 (~0.081 mg ml^{-1}). [θ] is the mean residue molar ellipticity. **c**, Transmission electron micrograph, obtained as described in Methods section, showing a network of polymer tapes. The tapes are ~7–8 nm wide and are entangled. Very few free ends are apparent, consistent with very long tapes. The density of the tapes decreased as the solution was diluted. With prolonged exposure, the structures degraded.

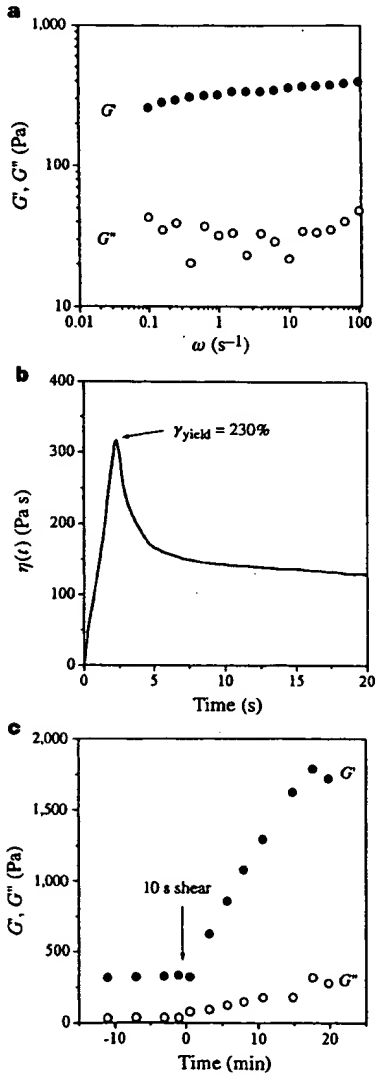

Figure 2 Viscoelastic properties of 0.019 volume fraction (~30 mg ml^{-1}) K24 in 2-chloroethanol at 25 °C. **a**, Typical mechanical spectra (elastic modulus G' and viscous modulus G'' versus frequency ω of applied strain) of fully set gels with 1% strain (see Methods), which is within the linear viscoelastic region. **b**, Plot of the stress growth $\eta(t)$ versus time obtained with steady shearing of the sample at a shear rate of 1 Hz. **c**, Growth of the plateau elastic and viscous moduli after shearing for 10 s, at a constant shear rate of 5 Hz. Data were collected with small oscillatory deformation experiments (5% strain), at a shear rate ω of 1 s^{-1}.

behaviour of the peptide Lysβ-21 whose structure (CH_3CO-Gln-Ala-Thr-Asn-Arg-Asn-Thr-Asp-Gly-Ser-Thr-Asp-Tyr-Gly-Ile-Leu-Gln-Ile-Asn-Ser-Arg-NH_2) corresponds to residues 41–61 of hen egg white lysozyme and forms a triple-stranded β-sheet in the β-domain of the native protein[15]. The first strand of this domain is exposed to water. The 'phase diagram' (Fig. 3b) is very similar to the one for K24, except that gels are now obtained generally in solvents with $\epsilon_r > 30$, including, for example, water ($\epsilon_r = 78$) and formamide ($\epsilon_r = 109$). The IR spectrum of the gel is consistent with a predominantly β-sheet structure.

Criteria for the rational design of gel-forming peptides have been deduced from the above results and information in the relevant literature[6,16–19]. In addition to the highly cooperative intermolecular hydrogen bonds, which are of primary importance for the generation of the tapes, these are: (1) cross-strand attractive forces (hydrophobic, electrostatic, hydrogen-bonding) between side-chains, (2) lateral recognition between adjacent β-strands to constrain their self-assembly to one dimension, and (3) strong adhesion of solvent to the surface of the tapes to control solubility. Based on these criteria, a *de novo* 11-residue peptide DN1, CH_3CO-Gln-Gln-Arg-Phe-Gln-Trp-Gln-Phe-Glu-Gln-Gln-NH_2, was designed to form β-sheet polymer tapes in water. Indeed, the peptide adopts a β-strand configuration in D_2O, as revealed by its IR spectrum (the major component of the amide I′ band is centred at 1,619 cm^{-1} and has half-height bandwidth of 20 cm^{-1}), and produces a self-supporting, thermostable gel, up to at least 90 °C, and at volume fractions of 0.01 (~15 mg ml^{-1}) or above at neutral pH in water (see Methods for details). The hydrophobic interactions are provided by the $(-CH_2-)_2$ moieties of the six glutamine residues (Fig. 4). The residues Phe 4, Trp 6 and Phe 8 are also hydrophobic, but they are also expected to provide intermolecular recognition by $\pi-\pi$ interactions[18,20]. Arg 3 and Glu 9 provide an additional degree of recognition via their strong coulombic attraction[18], and favour an antiparallel alignment of the strands. Gln, Arg and Glu side-chains make the β-sheet surface hydrophilic. Chemically blocked termini were used to avoid edge-to-edge coulombic attractions between tapes. In contrast, an 11-residue peptide CH_3CO-Gln-Gln-Gln-Gln-Gln-Gln-Gln-Gln-Gln-Gln-Gln-NH_2 assembles into β-sheet structures which are highly insoluble in water, consistent with the behaviour of polyglutamine-based proteins[21]. We believe that this behaviour stems from the absence of lateral recognition between β-strands and/or strong intersheet interactions as compared with those to the solvent.

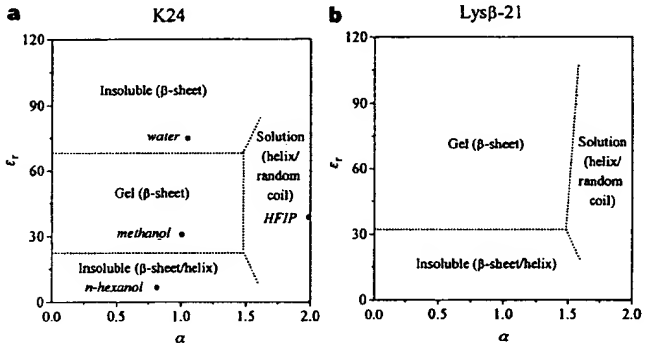

Figure 3 Conformational propensities and gelation properties of the peptides K24 (**a**) and Lysβ-21 (**b**), as a function of the relative dielectric constant ϵ_r and the hydrogen-bond donor ability α of the solvent (peptide volume fraction ~0.006, corresponding to ~10 mg ml^{-1}). The α-values for different solvents were obtained by the solvatochromic comparison method[13]. Details of the solvents used are given in the Methods section. The areas separated by the dotted lines represent distinctive behavioural patterns. The four points shown for K24 in **a** are examples of representative solvents lying within each region. Preliminary results with the *de novo* peptide DN1 show a similar phase diagram to that of Lysβ-21 (**b**), except that the gel band for the *de novo* peptide extends from $\epsilon_r \approx 40$ to $\epsilon_r \approx 90$, and the peptide becomes insoluble at $\epsilon_r > 90$.

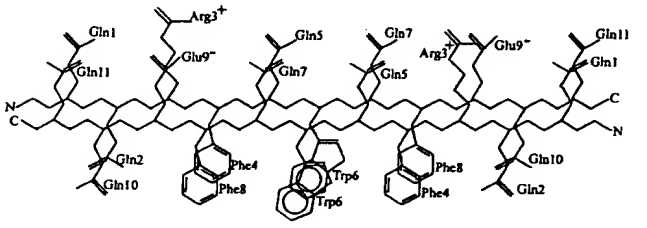

Figure 4 Proposed antiparallel β-sheet arrangement of two molecules of the *de novo* peptide DN1. The peptide backbones are drawn as black zig-zag lines. The N and C termini are indicated with the letters N and C, respectively. Side chains that carry a net charge at neutral pH are labelled with + or −.

An exploitable property of the peptide gels is their responsiveness to chemical and physical triggers. For example, the Lysβ-21 gels are stable up to pH 12, but are transformed into newtonian fluids at higher values, presumably as a consequence of deprotonation of the side-chains Arg 5 and Arg 21. Indeed, it is well known that peptide conformations can be switched by changes in the pH of the solution[6,15,22]. A gel-to-fluid transition can also be brought about by varying the hydrogen-bond donor strength of the solvent[13,23]. This can be achieved, for example, in the case of the K24/methanol system, by adding HFIP (see Figs 1b and 3a).

The novel rheological response of the K24 gel in 2-chloroethanol to a strong shear flow is an interesting example of a physical trigger (Fig. 2c). The elastic and viscous moduli are 'switched' over a timescale of 20 minutes into a state with a much higher modulus (by a factor of 5), but a much smaller linear range of strain (~2%, that is, a more brittle gel). Relaxation to the equilibrium state takes of the order of 10 hours. This unusual behaviour may be connected with annealing of flow-induced structural defects in the tapes. Such a phenomenon is not observed in linear polymers and other self-assembled polymers such as the cylindrical worm-like micelles formed by surfactants[24].

The β-sheet peptide tapes show some similarity, in structural terms, to protein fibrils formed from a variety of different proteins, including the polyglutamine-containing proteins responsible for Huntington's disease[21] and the more complex structure of amyloid fibres[10]. Other linear biological peptides, such as leucine-rich (LRR) fragments of the *Drosophila* Toll protein[25], a 28-oligonucleotide fragment of the β-amyloid protein[26] and peptides modelled on conserved domains of desmin[27], as well as synthetic peptides incorporating non-natural chemical groups[28–30] have been reported to form gels.

The peptide gels described here are potentially biocompatible and biodegradable and in this respect can be compared with biopolymer gels[31] such as gelatin, actin, amylose and agarose. The rheological measurements on K24 indicate comparable mechanical properties (elastic and viscous moduli) to these 'classic' biopolymer gels under linear deformation, but show much greater recoverable strain. The mechanically induced strengthening of the peptide gels could be used to programme gel properties simply by shearing. Thus, we believe that the polymeric β-sheet tapes will find a wide range of applications, and will also enable fundamental issues to be explored. Particular issues are the elucidation of the molecular mechanism and control of β-sheet folding pathways in both functional native proteins and those responsible for amyloidogenesis in neurological diseases, as well as the physics of tape-like polymers[32]. □

Methods

IR. Spectra were averages of four scans, recorded with a resolution of 4 cm^{-1}, at 20 °C in a 50-μm CaF_2 cell, using a Perkin Elmer 1760X FTIR spectrometer.

Component peaks were obtained by second-derivative analysis and band-fitting of the absorption spectra[33], after subtraction of the solvent spectrum.

CD. Measurements were performed at 20 °C on a Jasco J-715 spectropolarimeter using quartz cuvettes (path length 1 mm) and a response time of 1 s. Each spectrum was the average of four scans. The peptide concentration in solution was determined by amino acid analysis and ninhydrin assay. CD curves were fitted using the algorithm CCA (ref. 15).

TEM. The specimens were prepared by deposition of the peptide gel (0.001 volume fraction or 1.5 mg ml^{-1} in methanol, prepared 24 h before study to ensure complete formation of the gel and then diluted 20 times before use) onto a 300 mesh size, glow-discharged, carbon-coated, copper grid followed by negative staining with uranyl acetate solution (4% w/v in water). The samples were examined with a Phillips CM10 TEM at 100 kV accelerating voltage.

Rheology. A Rheometrics Dynamic Analyser II, with parallel plate (25 mm diameter) geometry, was employed. The thickness of the material between the plates was between 0.5 and 0.8 mm. Rubber elasticity theory[11,12] has been used to extract information about the structure and dynamic properties of the gel network. For example, the average distance ξ between nearest entanglements in space (mesh size) is: $\xi = \{(g_N f k_B T)/(2G^0_N)\}^{1/3}$ where G^0_N is the elastic modulus in the plateau region, measured from the mechanical spectra in the linear viscoelastic region, $g_N \approx 1$, k_B is Boltzmann's constant, T the absolute temperature, and f, the number of network tapes associated with each entanglement, is taken to have a value ≥ 4. The stress growth $\eta(t)$ of the material is monitored as a function of time, at constant shear rate $\dot{\gamma}$. The stress builds up proportionally to the duration of the applied shear. The strain at which the tapes are fully extended is $\gamma_{yield} \approx N_c^{1/2}$, where N_c is the number of persistence lengths of the tapes between nearest entanglements on the same tape. Further increase of the strain causes the network to break down and the stress to relax. Using the above equations for ξ and γ_{yield}, a lower limit can be placed on the persistence length $l \geq \xi/N_c^{1/2}$. The thickness, h, of a biaxial tape-like polymer can also be deduced, assuming that all tapes in solution participate in the network: $h = \phi_p g_N k_B T l / G^0_N N_c l w$, where w is the width of the tape and ϕ_p its volume fraction.

Peptide synthesis and purification. Peptides (K24[9], Lysβ-21[15] and the 11-mer *de novo* peptides) were synthesized and purified according to standard methods. K24 corresponds to residues 41–67 of IsK[9], with residues 47–49 missing, whilst K27 corresponds to residues 41–67 of IsK.

Phase diagrams. These were constructed from observations of the IR spectra and mechanical properties of solutions of the peptides in a variety of solvents. K24 in HFIP or in the mixtures 50%HFIP/50%CH_2Cl_2 or 90%HFIP/10% methanol forms clear fluid solutions with at least 75% of the peptide in the helical conformation, K24 solutions in solvents such as methanol, 2-chloroethanol, glycerol, ethylene glycol, dimethylformamide (DMF), 70%DMF/30%formamide, DMF/methanol(0–100%), 90%propanol/10%formamide, 50%propanol/50%formamide, water/propanol(30–90%) and HFIP/methanol (70–90%), lie in the gel region (β-sheet conformation ≥70%). K24 solutions in formamide, water, mixtures of 70%formamide/30%DMF, water/propanol (10–20%), 20%methanol/80%water, 50%HFIP/50%water, and 50% 2,2,2-trifluoroethanol (TFE) 50% water lie above the gel region, whereas solutions in acetone, tetrahydrofuran, diethylether, hexane, dichloromethane, chloroform, 50%methanol/50%dichloromethane, propanol, butanol, hexanol and decanol lie below the gel region. Lysβ-21 was studied in similar solvents. In the case of Lysβ-21, gelation is achieved initially only at elevated temperatures, depending on the solvent (55 °C for water at pH < 12). This is believed to be governed by the effect of temperature on the aggregation behaviour of the β-tapes. However, the gel, once formed, is stable over the range 10–80 °C, at pH < 12.

Received 28 October 1996; accepted 4 February 1997.

1. Lehn, J.-M. *Angew. Chem. Int. Edn Engl.* **29**, 1304–1319 (1990).
2. Ball, P. *Nature* **371**, 202–203 (1994).
3. Lokey, R. S. & Iverson, B. L. *Nature* **375**, 303–305 (1995).
4. Muller, A. & Beugholt, C. *Nature* **383**, 296–297 (1996).
5. Berg, R. H., Hvilsted, S. & Ramanujam, P. S. *Nature* **383**, 505–508 (1996).
6. Choo, D. W., Schneider, J. P., Graciani, N. R. & Kelly, J. W. *Macromolecules* **29**, 355–366 (1996).
7. Krejchi, M. T. *et al. Science* **265**, 1427–1432 (1994).
8. Takumi, T. *News Physiol. Sci.* **8**, 175–178 (1993).
9. Aggeli, A. *et al. Biochemistry* **35**, 16213–16221 (1996).
10. Blake, C. & Serpell, L. *Structure* **4**, 989–998 (1996).
11. Doi, M. & Edwards, S. F. *The Theory of Polymer Dynamics* (Clarendon, Oxford, 1986).
12. Ferry, J. D. *Viscoelastic Properties of Polymers* 2nd edn (Wiley, New York, 1970).
13. Kamlet, M. J., Abboud, J. L. M., Abraham, M. H. & Taft, R. W. *J. Org. Chem.* **48**, 2877–2887 (1983).
14. Mayo, K. H., Ilyina, E. & Park, H. *Protein Science* **5**, 1301–1315 (1996).
15. Yang, J. J., Pitkeathly, M. & Radford, S. E. *Biochemistry* **33**, 7345–7353 (1994).
16. Minor, D. L. & Kim, P. *Nature* **380**, 730–734 (1996).
17. Otzen, D. E. & Fersht, A. R. *Biochemistry* **34**, 5718–5724 (1995).
18. Smith, K. C. & Regan, L. *Science* **270**, 980–982 (1995).
19. Zhang, S., Holmes, T., Lockshin, C. & Rich, A. *Proc. Natl Acad. Sci. USA* **90**, 3334–3338 (1993).
20. Fasman, G. D. *Prediction of Protein Structure and the Principles of Protein Conformation* (Plenum, New York, 1989).
21. Stott, K., Blackburn, J. M., Butler, P. J. G. & Perutz, M. *Proc. Natl Acad. Sci. USA* **92**, 6509–6513 (1995).
22. Chirgadze, Yu. N., Shestopalov, B. V. & Venyaminov, S. Yu. *Biopolymers* **12**, 1337–1351 (1973).
23. Guo, H. & Karplus, M. *J. Phys. Chem.* **98**, 7104–7105 (1994).
24. Clausen, T. M. *et al. J. Phys. Chem.* **96**, 474–484 (1992).
25. Gay, N. J., Packman, L. C., Weldon, M. A. & Barna, J. C. J. *FEBS Lett.* **291**, 87–91 (1991).
26. Kirschner, D. A. *et al. Proc. Natl Acad. Sci. USA* **84**, 6953–6957 (1987).
27. Geisler, N., Heimburg, T., Schuneman, J. & Weber, K. *J. Struct. Biol.* **110**, 205–214 (1993).
28. Hanabusa, K., Naka, Y., Koyama, T. & Shirai, H. *J. Chem. Soc., Chem. Commun.* 2683–2684 (1994).
29. Stock, H. T., Turner, N. J. & McCague, R. *J. Chem. Soc., Chem. Commun.* 2063–2064 (1995).
30. Vegners, R., Shestakova, I., Kalvinsh, I., Ezzell, R. M. & Janmey, P. A. *J. Pept. Sci.* **1**, 371–378 (1995).
31. Clark, A. H. & Ross-Murphy, S. B. *Adv. Polym. Sci.* **83**, 57–192 (1987).
32. Nyrkova, I. A., Semenov, A. M., Joanny, J. F. & Khokhlov, A. R. *J. Phys. II* **6**, 1411–1428 (1996).
33. Surewicz, W. K. & Mantsch, H. H. *Biochim. Biophys. Acta* **952**, 115–130 (1988).

Acknowledgements. We thank A. Marsh, I. Nyrkova and A. N. Semenov for discussions on the physics of tape-like polymers; P. McPhie for assistance with the electron microscope; G. Mitchell for helping with the X-ray diffraction measurements at the Daresbury Laboratory; A. Smith and R. Owens for help with the atomic force microscopy experiments; and J. L. Johnson for assistance with peptide synthesis. This work was supported by the UK Engineering and Physical Sciences Research Council, the Wellcome Trust, and Schlumberger Cambridge Research. Lysβ-21 was synthesized at OCMS which is supported by the BBSRC, MRC and EPSRC.

Correspondence should be addressed to N.B. (e-mail: nevb@chem.leeds.ac.uk).

Low-latitude glaciation in the Palaeoproterozoic era

D. A. Evans*, N. J. Beukes† & J. L. Kirschvink*

* *Division of Geological and Planetary Sciences 170-25, California Institute of Technology, Pasadena, California 91125, USA*

† *Department of Geology, Rand Afrikaans University, Auckland Park 2006, Johannesburg, South Africa*

One of the most fundamental enigmas of the Earth's palaeoclimate concerns the temporal and spatial distributions of Precambrian glaciations. Through four billion years of Precambrian history, unequivocally glacial deposits have been found only in the Palaeoproterozoic and Neoproterozoic record[1]. Nonetheless, some of these deposits are closely associated with tropical—rather than just polar—palaeolatitudinal indicators such as carbonate rocks, red beds, and evaporites[1,2]. These observations are quantitatively supported by palaeomagnetic results indicating a ~5° latitude for Neoproterozoic glaciogenic rocks in Australia[3–5]. Similarly reliable palaeolatitudes for the older, Palaeoproterozoic glaciogenic rocks have not yet been obtained, as such deposits commonly suffer from poor preservation and secondary magnetic overprinting. The Archaean–Palaeoproterozoic 'Transvaal Supergroup' on the Kaapvaal craton in South Africa is, however, exceptionally well preserved, and is thus amenable to the palaeomagnetic determination of depositional palaeolatitudes. Within this supergroup the ~2.2 billion-year old Ongeluk lavas are a regionally extensive, largely undeformed and unmetamorphosed, extrusive volcanic succession[6], which conformably overlies glaciogenic deposits (the Makganyene diamictite). Here we report a palaeomagnetic estimate of 11 ± 5° depositional latitude for the lavas, and hence for the underlying contemporaneous glacial rocks. The palaeoclimate enigma is thus deepened; a largely ice-free Precambrian world was apparently punctuated by two long ice ages, both yielding glacial deposits well within tropical latitudes.

The Ongeluk Formation comprises ~500–1,000 m of extrusive, basaltic-andesitic lavas deposited in the marine facies succession of

IMMUNO BIOLOGY

THE IMMUNE SYSTEM IN HEALTH AND DISEASE

FOURTH EDITION

Charles A. Janeway, Jr.
Yale University Medical School

Paul Travers
Anthony Nolan Research Institute, London

Mark Walport
Imperial College School of Medicine, London

J. Donald Capra
Oklahoma Medical Research Foundation

CB
CURRENT BIOLOGY PUBLICATIONS

GARLAND PUBLISHING
Taylor & Francis Group

Text Editors:	Penelope Austin, Eleanor Lawrence
Editorial Assistant:	Richard Woof
Copyeditor:	Bruce Goatly
Production Editor:	Emma Hunt
Indexer:	Liza Weinkove
Illustration and Layout:	Blink Studio, London

Distributors:

Inside North America: Garland Publishing, 19 Union Square West, New York, NY 10003, US.
Inside Japan: Nankodo Co. Ltd., 42-6, Hongo 3-Chome, Bunkyo-ku, Tokyo 113, Japan.
Outside North America and Japan: Churchill Livingstone, Robert Stevenson House, 1–3 Baxter's Place, Leith Walk, Edinburgh, EH1 3AF.

ISBN 0 8153 3217 3 (paperback) Garland
ISBN 0 4430 6275 7 (paperback) Churchill Livingstone
ISBN 0 4430 6274 9 (paperback) International Student Edition

A catalog record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Janeway, Charles.
Immunobiology: the immune system in health and disease / Charles A. Janeway, Jr., Paul Travers, Mark Walport; with the assistance of J. Donald Capra.—4th ed.
p.cm.
ISBN 0-8153-3217-3 (pbk.)
1. Immunity.I. Travers, Paul, 1956–II. Walport, Mark.
III. Title.
QR181.J37 1999
616.07'9- - dc21

98-30316
CIP

This book was produced using QuarkXpress 3.32 and Adobe Illustrator 7.0.

Printed in United States of America.

Published by Current Biology Publications, part of Elsevier Science London
Middlesex House, 34–42 Cleveland Street, London W1P 6LB, UK
and Garland Publishing, a member of the Taylor & Francis Group,
19 Union Square West, New York, NY 10003, US.

distinguish between foreign material coming from the cytosolic and vesicular compartments. This is achieved through the delivery of peptides to the cell surface from each of these intracellular compartments by a different class of MHC molecule. MHC class I molecules deliver peptides originating in the cytosol to the cell surface, where they are recognized by CD8 T cells. MHC class II molecules deliver peptides originating in the vesicular system to the cell surface, where they are recognized by CD4 T cells (see Fig. 4.2). We shall see later, when we discuss the recognition of MHC molecules by the T-cell receptor, how the molecules CD8 and CD4 help in the differential recognition of MHC class I and MHC class II molecules by the two major subsets of T cells.

4-3 The two classes of MHC molecule have distinct subunit structures but similar three-dimensional structures.

The MHC class I and MHC class II molecules are cell-surface glycoproteins closely related in overall structure and function, although they have different subunit structures. Both molecules have two domains that resemble immunoglobulin domains, and two domains that fold together to create a long cleft that is the site where peptides bind. However, differences in their structures allow them to serve distinct functions in antigen presentation,

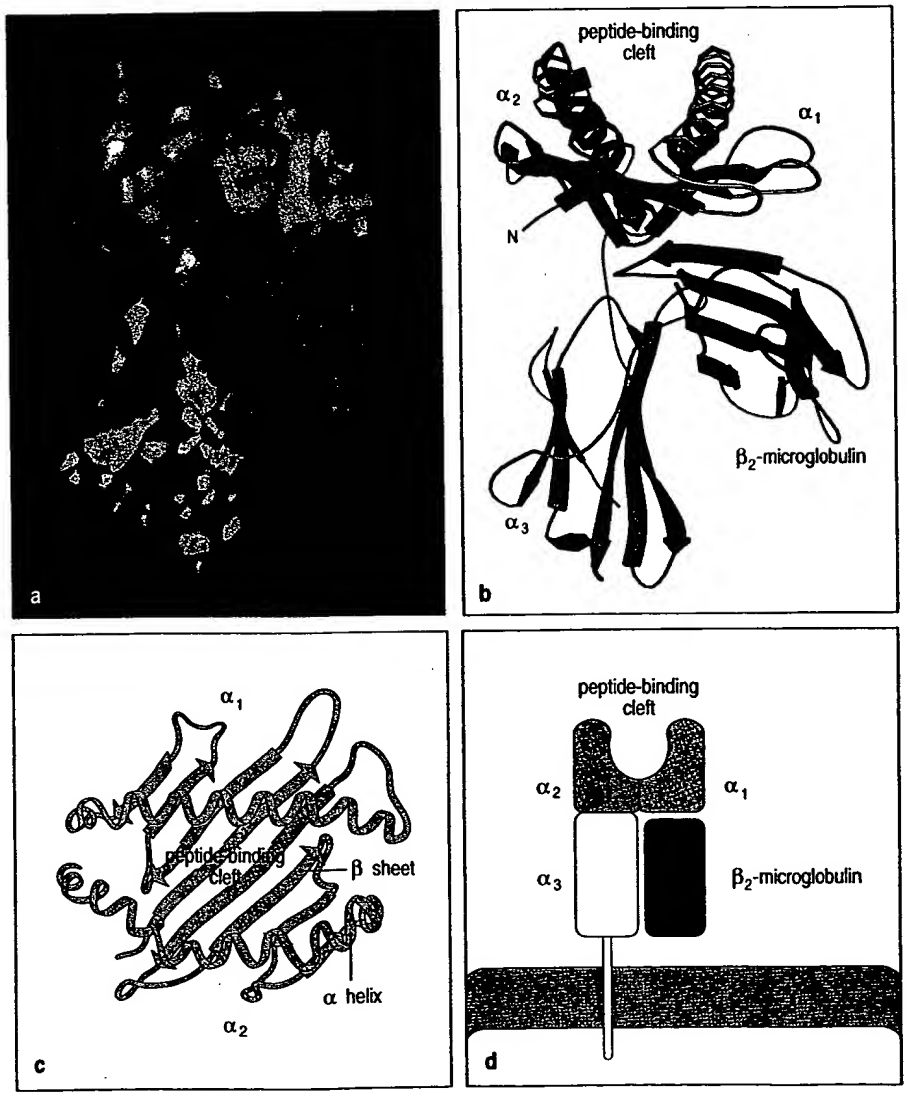

Fig. 4.3 The structure of an MHC class I molecule, determined by X-ray crystallography. Panel a shows a computer graphic representation of a human MHC class I molecule, HLA-A2, which has been cleaved from the cell surface by the enzyme papain. The surface of the molecule is shown, colored according to the domains described below. Panel b shows a ribbon diagram of that structure. Shown schematically in panel d, the MHC class I molecule is a heterodimer of a membrane-spanning α chain (molecular weight 43,000 Da), non-covalently associated with β_2-microglobulin (12,000 Da), which does not span the membrane. The α chain folds into three domains: α_1, α_2, and α_3. The α_3 domain and β_2-microglobulin show similarities in amino acid sequence to immunoglobulin constant domains and have similar folded structures, whereas the α_1 and α_2 domains fold together into a single structure consisting of two segmented α helices lying on a sheet of eight antiparallel β strands. The folding of the α_1 and α_2 domains creates a long cleft or groove, which is the site at which peptide antigens bind to the MHC molecules. The transmembrane region and the short stretch of peptide that connects the external domains to the cell surface are not seen in panels a and b as they have been removed by the papain digestion. As can be seen in panel c, looking down on the molecule from above, the sides of the cleft are formed from the inner faces of the two α helices; the β-pleated sheet formed by the pairing of the α_1 and α_2 domains creates the floor of the cleft. We shall use the schematic representation in panel d throughout this text.

binding peptides from different intracellular sites and activating different subsets of T cells. Purified peptide:MHC class I and peptide:MHC class II complexes have been characterized structurally, allowing us to describe in detail both the MHC molecules themselves and the way in which they bind peptides.

MHC class I structure is outlined in Fig. 4.3. MHC class I molecules consist of two polypeptide chains, an α or heavy chain encoded in the MHC, and a smaller non-covalently associated chain, β_2-microglobulin, which is not encoded in the MHC. Only the class I α chain spans the membrane. The molecule has four domains, three formed from the MHC-encoded α chain, and one contributed by β_2-microglobulin. The α_3 domain and β_2-microglobulin have a folded structure that closely resembles that of an immunoglobulin domain (see Section 3-5). The most remarkable feature of MHC class I molecules is the structure of the α_1 and α_2 domains, which pair to generate a cleft on the surface of the molecule that is the site of peptide binding.

MHC class II molecules consist of a non-covalent complex of two chains, α and β, both of which span the membrane (Fig. 4.4). The crystal structure of the MHC class II molecule shows that it is folded very much like the MHC class I molecule. The major differences lie at the ends of the peptide-binding cleft, which are more open in MHC class II molecules. The main consequence

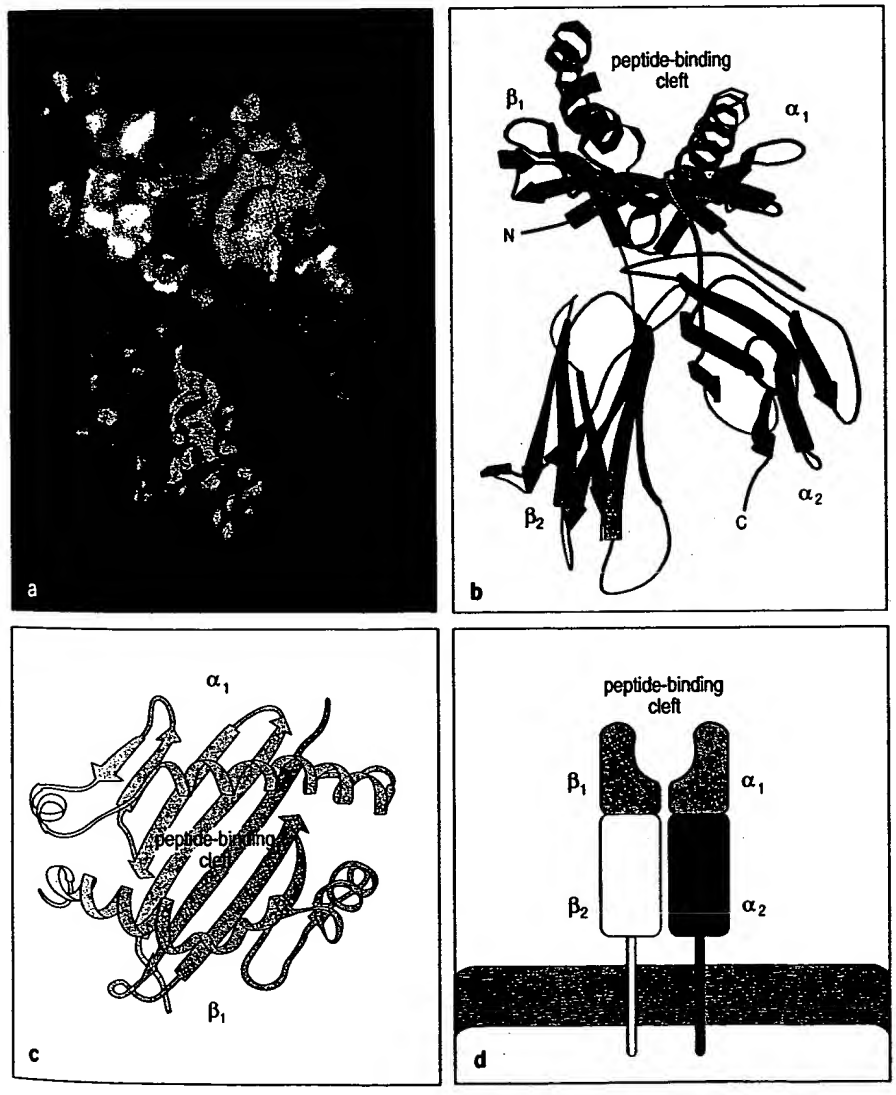

Fig. 4.4 MHC class II molecules resemble MHC class I molecules in structure. The MHC class II molecule is composed of two transmembrane glycoprotein chains, α (34,000 Da) and β (29,000 Da), as shown schematically in panel d. Each chain has two domains, and the two chains together form a compact four-domain structure similar to that of the MHC class I molecule (compare with panel d of Fig. 4.3). Panel a shows a computer graphic representation of the surface of the MHC class II molecule, in this case the human protein HLA-DR1, and panel b shows the equivalent ribbon diagram. The α_2 and β_2 domains, like the α_3 and β_2-microglobulin domains of the MHC class I molecule, have amino acid sequence and structural similarities to immunoglobulin constant domains; in the MHC class II molecule, the two domains forming the peptide-binding cleft are contributed by different chains and are therefore not joined by a covalent bond (see panels c and d). Another important difference, not apparent in this diagram, is that the peptide-binding groove of the MHC class II molecule is open at both ends.